



TAÍS RODRIGUES DA COSTA

**RISCO DE INSOLVÊNCIA: UMA ANÁLISE DE DIFERENTES
MÉTODOS DE PREVISÃO E DO EFEITO DA ESTRUTURA DE
CAPITAL**

**LAVRAS - MG
2023**

TAÍS RODRIGUES DA COSTA

**RISCO DE INSOLVÊNCIA: UMA ANÁLISE DE DIFERENTES MÉTODOS DE
PREVISÃO E DO EFEITO DA ESTRUTURA DE CAPITAL**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Administração, área de concentração em Estratégia de Negócios Globais e Finanças Corporativas, para a obtenção do título de Mestre.

Prof. Dr. José Willer do Prado
Orientador

**Lavras - MG
2023**

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Costa, Taís Rodrigues.

Risco de insolvência: Uma análise de diferentes métodos de
previsão e do efeito da estrutura de capital / Taís Rodrigues Costa. -
2023.

169 p.

Orientador(a): José Willer do Prado.

Dissertação (mestrado acadêmico) - Universidade Federal de
Lavras, 2023.

Bibliografia.

1. Determinantes do risco de insolvência. 2. Modelos
preditivos. 3. Insolvência empresarial. I. do Prado, José Willer. II.
Título.

TAÍS RODRIGUES DA COSTA

**RISCO DE INSOLVÊNCIA: UMA ANÁLISE DE DIFERENTES MÉTODOS DE
PREVISÃO E DO EFEITO DA ESTRUTURA DE CAPITAL**

**INSOLVENCY RISK: AN ANALYSIS OF DIFFERENT FORECASTING METHODS
AND THE EFFECT OF CAPITAL STRUCTURE**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós- Graduação em Administração, área de concentração em Estratégia de Negócios Globais e Finanças Corporativas, para a obtenção do título de Mestre.

APROVADO em 24 de março de 2023.

Prof. Dr. Ednilson Sebastião de Ávila

UFLA

Prof. Dr. Lélis Pedro de Andrade

IFMG

Prof. Dr. Paulo Henrique Sales Guimarães

UFLA

Prof. Dr. José Willer do Prado
Orientador

**Lavras - MG
2023**

Dedico aos meus pais, Jandir e Romilda, e ao meu irmão, Tiago.

AGRADECIMENTOS

Agradeço, primeiramente a Deus, por me conceder vida, saúde e força de vontade para ir em busca dos meus objetivos, bem como por me proporcionar boas oportunidades de crescimento e amadurecimento. Meu imenso agradecimento aos meus pais, Jandir e Romilda, por todo cuidado, amor, incentivo e por serem a minha base por todos os anos da minha vida. Agradeço ao meu irmão Tiago, por todo apoio, paciência, amor e companheirismo.

Agradeço imensamente ao meu orientador, José Willer do Prado, por toda ajuda, incentivo e conhecimento compartilhado durante a orientação do mestrado e co-orientação do trabalho de conclusão de curso. Agradeço também à minha orientadora do trabalho de conclusão de curso da graduação em Administração, Adriana Giarola Vilamaior, por todo carinho, incentivo e conhecimento compartilhado durante a orientação, que foram decisivos para o meu ingresso na pós-graduação. Agradeço aos professores Ednilson Sebastião de Ávila, Lélis Pedro de Andrade e Paulo Henrique Sales Guimarães, por terem aceitado o convite para serem membros das bancas de qualificação e defesa, e por todas as considerações realizadas, bem como pela ajuda durante o desenvolvimento do estudo.

Agradeço à Universidade Federal de Lavras e a todos os servidores, em especial aos professores do Programa de Pós-Graduação em Administração (PPGA-UFLA), pelos ensinamentos durante o período do mestrado. Agradeço também aos colegas do PPGA, especialmente, Fabiane Fidelis Querino e Lara Carvalho Vilela, pela amizade, apoio e troca de conhecimentos e experiências. Meu agradecimento também ao órgão de fomento: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Agradeço também minha tia Sudária Maria da Costa, por toda ajuda e incentivo, em especial durante os anos de graduação. Por fim, agradeço aos verdadeiros amigos, que se fizeram presentes e me apoiaram durante neste e outros períodos da minha vida, especialmente minhas amigas Juliana, Karla e Bruna.

*“O insucesso é apenas uma oportunidade para recomeçar com
mais inteligência.” (Henry Ford)*

RESUMO

O objetivo geral do estudo foi analisar o risco de insolvência de empresas brasileiras não financeiras listadas na B3, sob as perspectivas da previsão de insolvência e estrutura de capital, analisando a assertividade de modelos de previsão durante o momento de pandemia e os efeitos da COVID-19 na influência da estrutura de capital sobre o risco de insolvência. Os objetivos específicos foram: Realizar uma revisão bibliométrica sobre previsão e risco de insolvência (artigo 1); identificar entre diferentes técnicas de previsão qual fornece resultados mais assertivos para o período de pandemia (artigo 2); analisar o efeito da estrutura de capital no risco de insolvência e o impacto da pandemia como moderadora nessa relação (artigo 3). Foram usadas as técnicas de bibliometria (artigo 1), análise discriminante, regressão logística, k-vizinhos mais próximos, árvore de decisão, floresta aleatória, redes neurais artificiais e máquinas de vetores de suporte (artigo 2), e regressão multinível de dados em painel (artigo 3). Como resultado, o artigo 1 mostrou que o campo de estudos se desenvolveu lentamente até os anos 2000, com tendência de crescimento a partir de 2008. Posteriormente, houve aumentos consideráveis a partir do ano de 2016, com auge de publicações em 2022. O estudo mais relevante foi o de Altman (1968). O principal país para o campo de estudos foi os Estados Unidos, mas a China apresentou o maior volume de publicações. O periódico mais relevante foi o *Expert Systems With Application*, e ao analisar as palavras-chave do campo de estudo, as principais foram *classification*, *financial ratios*, *risk*, *neural-networks* e *models*. Os tópicos de tendência foram computacionais como *big data*, *machine learning* e *deep learning*. No artigo 2, a técnica com melhor percentual de acurácia foi a de árvore de decisão, seguida pela de floresta aleatória, regressão logística e análise discriminante, rede neural artificial, k-vizinhos mais próximos e máquina de vetores de suporte. Tal resultado mostrou que as técnicas de caixa branca, caixa cinza e tradicionais tiveram melhores acurácias em relação às técnicas de caixa preta (RNA e SVM). Por meio do artigo 3, foi possível identificar que algumas das variáveis de estrutura de capital apresentaram resultados significativos e positivos, confirmando a influência positiva do endividamento no risco de insolvência. A variável endividamento financeiro e seu termo quadrático apresentaram resultados significativos, que confirmaram a relação em forma de U entre o endividamento financeiro e o risco de insolvência. Em relação ao efeito moderador da pandemia, apenas a hipótese de que a pandemia de COVID-19 fortaleceu a influência positiva do endividamento de longo prazo no risco de insolvência foi confirmada. Este estudo contribuiu para a literatura de risco e previsão de insolvência ao realizar uma análise bibliométrica ampla, considerando duas bases de dados relevantes (*Web of Science* e *Scopus*); ao desenvolver modelos de previsão de insolvência considerando diferentes técnicas de previsão e analisando a assertividade para o momento de pandemia; e analisando a influência da estrutura de capital no risco de insolvência, bem como o efeito moderador da pandemia.

Palavras-chave: Determinantes do risco de insolvência. Modelos preditivos. Insolvência empresarial.

ABSTRACT

The overall objective of the study was to analyze the insolvency risk of Brazilian non-financial companies listed on B3, from the perspective of insolvency forecasting and capital structure, analyzing the assertiveness of forecasting models during the pandemic moment and the effects of COVID-19 on the influence of capital structure on the risk of insolvency. The specific objectives were: To carry out a bibliometric review on prediction and risk of insolvency (article 1); identify, among different forecasting techniques, which provides more assertive results for the pandemic period (article 2); analyze the effect of capital structure on the risk of insolvency and the impact of the pandemic as a moderator in this relationship (Article 3). Bibliometric techniques (article 1), discriminant analysis, logistic regression, k-nearest neighbors, decision tree, random forest, artificial neural networks and support vector machines (article 2) and multilevel data regression were used in panel (article 3). As a result, article 1 showed that the field of study developed slowly until the 2000s, with a tendency to grow from 2008 onwards. most relevant was that of Altman (1968). The main country for the field of studies was the United States, but China had the highest volume of publications. The most relevant journal was Expert Systems With Application, and when analyzing the keywords of the field of study, the main ones were classification, financial ratios, risk, neural-networks and models. Trending topics were computational like big data, machine learning, and deep learning. In article 2, the technique with the best percentage of accuracy was decision tree, followed by random forest, logistic regression and discriminant analysis, artificial neural network, k-nearest neighbors and support vector machine. This result showed that the white box, gray box and traditional techniques had better accuracy in relation to the black box techniques (ANN and SVM). Through Article 3, it was possible to identify that some of the capital structure variables showed significant and positive results, confirming the positive influence of debt on the risk of insolvency. The financial indebtedness variable and its quadratic term showed significant results, which confirmed the U-shaped relationship between financial indebtedness and the risk of insolvency. Regarding the moderating effect of the pandemic, only the hypothesis that the COVID-19 pandemic strengthened the positive influence of long-term debt on the risk of insolvency was confirmed. This study contributed to the risk and insolvency prediction literature by performing a comprehensive bibliometric analysis, considering two relevant databases (Web of Science and Scopus); by developing insolvency prediction models considering different forecasting techniques and analyzing the assertiveness for the moment of a pandemic; and analyzing the influence of capital structure on insolvency risk, as well as the moderating effect of the pandemic.

Keywords: Determinants of the risk of insolvency. Predictive models. Business insolvency.

LISTA DE FIGURAS

Figura 1 — Modelo conceitual da dissertação.	25
Figura 2 — Teoria do <i>trade-off</i>	29
Figura 3 — Classificações da pesquisa por categoria.	33
Figura 4 — Resumo da proposta de metodologia para a dissertação.	36
Figura 5 — <i>String</i> final <i>Web of science</i>	58
Figura 6 — <i>String</i> final Scopus.	58
Figura 7 — Tendências temporais.	60
Figura 8 — Referência publicação ano espectroscopia.	63
Figura 9 — Evolução do volume de publicações por país.	64
Figura 10 — Mapa de colaboração entre os países e países do autor correspondente.	65
Figura 11 — Periódicos com maiores volumes de publicações.	66
Figura 12 — Periódicos mais citados.	67
Figura 13 — Autores mais citados.	67
Figura 14 — Gráfico de três campos: volume de publicações por autor.	68
Figura 15 — Análise de palavras-chave.	69
Figura 16 — Tópicos de tendência.	70
Figura 17 — Mapa temático do campo de estudos.	71
Figura 18 — Funcionamento da técnica KNN.	90
Figura 19 — Árvore de decisão para a classificação de animais.	92
Figura 20 — Representação de RNA com uma camada oculta.	94
Figura 21 — Processamento na rede neural artificial.	95
Figura 22 — Exemplo de aplicação do SVM.	97
Figura 23 — Curva ROC de todos os modelos de previsão.	114
Figura 24 — Interceptos aleatórios por setor.	149
Figura 25 — Interceptos aleatórios por empresa.	150
Figura 26 — Comparação MQO e HLM3.	152

LISTA DE GRÁFICOS

Gráfico 1 — Número de empresas insolventes por año.....	105
---	-----

LISTA DE QUADROS

Quadro 1 — Resumo de estudos sobre previsão de insolvência.	31
Quadro 2 — Estágios de elaboração do estudo.	56
Quadro 3 — Indicadores econômico-financeiros de estudos anteriores (continua).	106

LISTA DE TABELAS

Tabela 1 — Artigos com maiores volumes de citações.	61
Tabela 2 — Empresas em recuperação judicial e empresas correspondentes (continua).	101
Tabela 3 — Empresas em recuperação extrajudicial e empresas correspondentes.	102
Tabela 4 — Empresas com patrimônio líquido negativo e correspondentes (continua).	102
Tabela 5 — Indicadores econômico-financeiros usados nos modelos preditivos e VIF.	111
Tabela 6 — Estatísticas descritivas dos indicadores econômico-financeiros.	112
Tabela 7 — Resultados das técnicas de previsão.	112
Tabela 8 — Variáveis independentes do estudo.	139
Tabela 9 — Estatística descritiva.	145
Tabela 10 — Valores do fator de inflação da variância (continua).	145
Tabela 11 — Matriz de correlação.	146
Tabela 12 — Resultados do modelo nulo.	147
Tabela 13 — Modelo de tendência linear com interceptos aleatórios.	148
Tabela 14 — modelo de tendência linear com interceptos e inclinações aleatórias (continua).	150
Tabela 15 — Resultados do modelo completo.	153
Tabela 16 — Resultados do modelo completo incluindo as variáveis moderadoras para o período de pandemia (continua).	156

LISTA DE SIGLAS

ASS.	Assimetria.
AUC.	Área sob a curva (<i>area under the curve</i>).
CP.	Curto prazo.
CURT.	Curtose.
DIV.	Dívida.
D.PADRÃO.	Desvio padrão
DT.	Árvore de decisão (<i>decision tree</i>).
END.	Endividamento.
HLME3.	Modelo de regressão Multinível.
LDA.	Análise discriminante.
LIQ.	Líquido(a).
LP.	Longo prazo.
KNN.	K-vizinhos mais próximos.
LR-teste.	Teste de razão de verossimilhança.
MÁX.	Máximo.
MÉD.	Média.
MÍN.	Mínimo.
MQO.	Mínimos quadrados ordinários.
PL.	Patrimônio líquido.
RF.	Floresta aleatória (<i>random forest</i>).
RL.	Regressão logística.
ROC.	<i>Receiver operating characteristic curve</i> .
RNA.	Rede neural artificial.
SVM.	Máquina de vetores de suporte (<i>support vector machines</i>).
VARIÁV.	Variável.
VAR.	Variância.
VIF.	Fator de inflação da variância.

SUMÁRIO

	PRIMEIRA PARTE	17
1	INTRODUÇÃO	17
1.1	Contextualização	17
1.2	Problema de pesquisa	21
1.3	Objetivo geral	22
1.4	Objetivos específicos.....	22
1.5	Justificativas de pesquisa	22
1.6	Estrutura da dissertação	24
2	REFERENCIAL TEÓRICO	26
2.1	Insolvência empresarial.....	26
2.2	Insolvência de saldo e Estrutura de Capital	27
2.3	Risco de insolvência (ou risco de falência).....	30
3	METODOLOGIA GERAL	33
4	CONSIDERAÇÕES FINAIS	38
	REFERÊNCIAS	41
	SEGUNDA PARTE – ARTIGOS	46
	ARTIGO 1 – RISCO E PREVISÃO DE INSOLVÊNCIA: UMA REVISÃO BIBLIOMÉTRICA	47
1	INTRODUÇÃO	48
2	O QUE SE TEM PESQUISADO SOBRE INSOLVÊNCIA E RISCO DE CRÉDITO	50
3	METODOLOGIA	55
3.1	Estágio 1: operacionalização da pesquisa	56
3.2	Estágio 2: procedimentos de pesquisa	57
3.3	Estágio 3: Procedimentos de seleção	59
3.4	Estágio 4: análise da produção científica:	59
4	RESULTADOS	60
4.1	Tendências temporais e análises de volume de publicação	60
4.2	Análise de artigos mais citados	61
4.3	Análise do país de origem.....	64
4.4	Análise dos periódicos mais citados	66
4.5	Análises de autores Citados.....	67

4.6	Análise de palavras-chave	69
5	CONSIDERAÇÕES FINAIS	73
	REFERÊNCIAS BIBLIOGRÁFICAS	75
	ARTIGO 2 – PREVISÃO DE INSOLVÊNCIA: UMA ANÁLISE COMPARATIVA DE DIFERENTES TÉCNICAS DE PREVISÃO PARA O MOMENTO DE PANDEMIA	78
1	INTRODUÇÃO	79
2	REFERENCIAL TEÓRICO.....	82
2.1	Técnicas estatísticas para previsão	85
2.1.1	Análise discriminante	86
2.1.2	Regressão logística (<i>logit</i>)	87
2.2	Técnicas de aprendizado de máquina.....	88
2.2.1	K-vizinhos mais próximos (K-nearest neighbors – KNN).....	89
2.2.2	Árvore de decisão (<i>Decision tree</i> - DT)	90
2.2.3	Floresta aleatória (<i>Random Forest</i> - RF)	93
2.2.4	Redes neurais artificiais (RNAs).....	94
2.2.5	Máquina de vetores de suporte (<i>Support vector machines</i> -SVMs) -)	95
2.3	Curva ROC.....	97
3	METODOLOGIA	99
3.1	Procedimentos de coletados dados	99
3.2	Amostra	99
3.3	Variáveis - Indicadores Econômico-financeiros	106
3.4	Procedimentos de análise	109
4	RESULTADOS.....	111
5	CONSIDERAÇÕES FINAIS	115
	REFERÊNCIAS	118
	ARTIGO 3 – EFEITOS DA ESTRUTURA DE CAPITAL NO RISCO DE INSOLVÊNCIA.....	122
1	INTRODUÇÃO	123
2	REFERENCIAL TEÓRICO.....	126
2.1	Teoria do <i>Trade-off</i>	126
2.2	Estudos envolvendo a estrutura de capital	127
2.3	Estrutura de capital em mercados emergentes	130
2.4	Estrutura de capital e risco de insolvência em períodos de crise	130

2.5	Desenvolvimento de hipóteses sobre a influência da estrutura de capital no risco de falência.....	132
3	METODOLOGIA	137
3.1	Procedimentos de coleta de dados.....	137
3.2	Amostra	137
3.3	Variáveis	138
3.4	Procedimentos de análise	141
4	RESULTADOS.....	145
5	CONSIDERAÇÕES FINAIS	160
	REFERÊNCIAS	165

PRIMEIRA PARTE

1 INTRODUÇÃO

Nesta seção está a introdução, que irá proporcionar um entendimento inicial do estudo. Como componentes da introdução estão a contextualização do tema, em seguida está o problema de pesquisa, o objetivo geral do estudo, os objetivos específicos que serão desenvolvidos como artigos e a justificativa, que mostra a relevância do estudo. Por fim a estrutura do trabalho será detalhada.

1.1 Contextualização

O risco de insolvência se constitui como um importante fator para todos os que possuem alguma ligação com uma empresa que pode se tornar insolvente, como tomadores de decisão, fornecedores investidores e funcionários. O risco de insolvência pode ser entendido como a possibilidade que uma empresa se torne insolvente. Ou seja, o risco de que a empresa não seja capaz de gerar fluxos de caixa suficientes para a liquidação de suas dívidas dentro do prazo de vencimento (MATEUS, 2010). Tal fato pode culminar em situações como recuperação extrajudicial, recuperação judicial ou, em casos extremos, na liquidação (falência) da empresa.

O aumento do risco de insolvência e suas possíveis consequências geram custos adicionais para a empresa. Esses custos são chamados de custos da dificuldade financeira ou também podem ser genericamente chamados de custos de falência (JUPETIPE *et al.*, 2017). Os custos de dificuldades financeiras são os custos decorrentes da existência dessas dificuldades financeiras na empresa, podendo ser considerados como custos diretos ou indiretos (CASTRO JÚNIOR, 2003; JUPETIPE *et al.*, 2017).

Os custos diretos podem ser considerados como aqueles de fácil mensuração, como taxas pagas durante o processo de reestruturação e honorários de funcionários, sendo também geralmente os que possuem valores mais elevados. Já os custos indiretos são aqueles considerados mais difíceis de serem mensurados, podendo geralmente estar atrelados a oportunidades de negócios perdidas, devido às questões jurídicas que devem ser respeitadas durante os processos de recuperação e que podem limitar as tomadas de decisão (CASTRO JÚNIOR, 2003; JUPETIPE *et al.*, 2017).

De maneira complementar, Shi e Li (2019a) afirmaram que situações de insolvência podem levar empresas a encerrar suas atividades. Tal fato afeta diretamente todos aqueles que

possuem alguma ligação com a organização, podendo inclusive gerar um efeito sistêmico na economia (BATTISTON *et al.*, 2007). Sendo assim, a antecipação a problemas financeiros pode evitar que as empresas cheguem à situações indesejadas, como a de falência, fato que evita a ocorrência de custos diretos e indiretos da dificuldade financeira e protege os agentes que possuem ligação com a empresa.

Diante da importância do risco de insolvência, revisões sistemáticas sobre o tema têm sido realizadas pela literatura, para o mapeamento do campo de estudos (PRADO *et al.*, 2016; SHI; LI, 2019a, 2019b). Essas revisões, no entanto, geralmente consideram uma única base de dados, sendo elas *Web of Science* ou *Scopus*, para o mapeamento do campo de estudos. De acordo com Echchakoui (2020), estudos de revisões sistemáticas considerando essas bases de dados individualmente apresentam limitações, uma vez que pode-se entender que elas são complementares. Sendo assim, trabalhos atualizados sobre o campo de estudos envolvendo o risco de insolvência e considerando ambas as bases de dados (*Web of Science* e *Scopus*) são necessários para um mapeamento mais amplo do campo de estudos.

A previsão de insolvência consiste em uma importante ferramenta para o campo de estudos de risco de insolvência, uma vez que pode dar base às decisões de crédito e assim beneficiar os agentes envolvidos com a organização (PRADO *et al.*, 2016). Por meio de modelos preditivos é possível prever o evento de insolvência, principalmente por meio do uso de indicadores econômico-financeiros (SILVA *et al.*, 2018). Esses modelos tornam-se relevantes para os diversos agentes ligados à empresa, uma vez que a partir da antecipação da situação de insolvência a empresa pode definir estratégias para reverter o cenário.

Face aos possíveis desdobramentos de situações de insolvência, estudiosos têm desenvolvido modelos preditivos e analisado os efeitos do risco de falência para as empresas. No trabalho seminal de Altman (1968), o autor desenvolveu o primeiro modelo de previsão utilizando análise multivariada. A partir de então a pesquisa se desenvolveu e novas técnicas passaram a ser usadas, como técnicas de aprendizado de máquina (PRADO *et al.*, 2016). O estudo de Odom e Sharda (1990) introduziu o uso da técnica de redes neurais artificiais para determinação de um modelo de previsão de insolvência.

No Brasil, alguns estudos de previsão de insolvência podem ser mencionados, como Kanitz (1974), Elizabetsky (1976), Altman, Baidya e Dias (1979), Silva (1982), Sanvicente e Minardi (1998), Stüpp (2015), Prado *et al.* (2020), Pinto (2021) e Santos (2021). Os estudos brasileiros se concentraram no desenvolvimento de modelos de previsão considerando técnicas tradicionais como a análise discriminante e a regressão logística.

Na maioria das vezes, os estudos brasileiros também não se concentraram na comparação entre diferentes técnicas (BARBOZA; KIMURA; ALTMAN, 2017; PINTO, 2021; SANTOS, 2021), uma vez que diferentes técnicas fornecem diferentes níveis de assertividade (BARBOZA; KIMURA; ALTMAN, 2017). Menciona-se também que diante de períodos em que se tem situações não previstas como a pandemia de COVID-19, o poder de preditivo de técnicas de previsão pode ser impactado (KITOWSKI; KOWAL-PAWUL; LICHOTA, 2022), sendo relevante, portanto, analisar se entre diferentes técnicas de previsão existe alguma que apresenta resultados mais robustos para uma amostra composta por empresas brasileiras, análise até então não encontrada na literatura.

Alguns estudos mostram que técnicas de aprendizado de máquina apresentam resultados mais robustos se comparadas com técnicas estatísticas (BARBOZA; KIMURA; ALTMAN, 2017; PINTO, 2021). Ademais, existem diferentes tipos de técnicas de aprendizado de máquina que são usadas para a elaboração de modelos preditivos. Essas técnicas podem ser classificadas em técnicas de caixa branca, caixa cinza e caixa preta, e geralmente se assume que as técnicas de caixa preta são as mais robustas (OBERMANN; WAACK, 2015; SJÖBERG *et al.*, 1995). No entanto, percebe-se que em alguns estudos as técnicas de caixa branca (OBERMANN; WAACK, 2015) ou de caixa cinza (BARBOZA; KIMURA; ALTMAN, 2017) apresentaram resultados superiores, o que torna necessário que novas investigações que comparem diferentes técnicas sejam realizadas.

De maneira complementar, Merton (1974) desenvolveu uma teoria de precificação de títulos considerando a estrutura de capital e o risco de falência. O trabalho de Merton (1974) se baseou no modelo de precificação de opções *Black-Scholes*, adaptando-o para a elaboração de um modelo que explicasse a precificação da dívida. De acordo com Merton (1974), o valor da dívida depende de três fatores principais, que são a taxa de retorno sobre a dívida sem risco que é exigida, as disposições e restrições do contrato e a probabilidade (risco) de inadimplência.

Ademais, as variáveis que influenciam o valor da dívida são o valor de mercado da empresa, o tempo (prazo de pagamento), a taxa de juros, a volatilidade do valor da empresa, a política de pagamento da empresa e a política de pagamento prometida aos credores (MERTON, 1974). Segundo o estudo de Merton (1974), a empresa entra em situação de falência quando o valor dos ativos (ou seja, o valor de mercado da empresa) cai e alcança um nível crítico. Isto é, a empresa se torna insolvente quando o seu valor é menor que o valor da dívida na data de pagamento (CHAIA, 2003; MERTON, 1974).

O risco de insolvência (ou probabilidade de inadimplência) pode ser decorrente de diversos fatores. A estrutura de capital da empresa é um fator considerado determinante para

esse risco, pois apresenta o nível de endividamento da empresa, o que impacta no risco e também no desempenho empresarial (TAO *et al.*, 2020). No entanto, existem visões distintas sobre os efeitos da estrutura de capital no valor e no risco das empresas (BRITO; CORRAR; BATISTELLA, 2007).

Assim como a elaboração de modelos preditivos, a identificação de fatores que influenciam no risco de insolvência também se faz necessária. No entanto, a literatura do campo se voltou principalmente para os estudos de previsão, sendo necessário que estudos sobre os fatores que influenciam nesse risco também sejam necessários. Conforme mencionado, a estrutura de capital é um fator de influência para o risco de insolvência, mas existem diferentes visões sobre os efeitos da estrutura de capital no valor e no risco das empresas (BRITO; CORRAR; BATISTELLA, 2007), existindo também resultados mistos no que se refere à relação existente entre endividamento e risco de insolvência (LOTT, 2019), que geralmente também são menos analisados simultaneamente.

Diante disso e face às possíveis consequências do risco de insolvência, torna-se relevante entender quais fatores influenciadores. A partir de tal conhecimento, principalmente os tomadores de decisão se beneficiarão, sendo possível formular estratégias para diminuir o risco de insolvência e evitar suas consequências. Sendo assim, a ampliação do conhecimento sobre risco de insolvência por meio da análise da influência da estrutura de capital, se faz necessária.

A existência de eventos não previstos, no entanto, é um fator que aumenta o risco de insolvência e dificulta sua previsão. Isso ocorre porque esses eventos trazem consequências imediatas, que antes de sua ocorrência não seria possível prever. Ou seja, as técnicas de previsão geralmente não são capazes de captar efeitos de eventos imprevisíveis (KITOWSKI; KOWAL-PAWUL; LICHOTA, 2022).

Um exemplo de evento não previsto que causou impacto na economia global foi a pandemia de COVID-19. Segundo o *The World Bank* (2021), a economia mundial sofreu uma retração de 4,3% em 2020. Apesar de a recuperação global ter sido retomada a partir de 2021 (THE WORLD BANK, 2022), as consequências da pandemia foram profundas. Houve um impacto em relação à mortes e doenças e a pandemia fez com que milhões de pessoas ficassem em situação de pobreza extrema.

A atividade econômica poderá demorar um longo período para se recuperar, uma vez que fatores como ameaças de variantes novas, aumento da inflação, aumento do endividamento e da desigualdade de renda podem dificultar a recuperação, principalmente de países emergentes e em desenvolvimento (THE WORLD BANK, 2021, 2022). Dessa maneira, os

efeitos causados pela pandemia continuam gerando incerteza e impactando na economia mundial, bem como na capacidade de geração de caixa das empresas. Sendo assim, o desempenho econômico-financeiro das empresas pode ser comprometido, elevando o risco de falência, fato comum em períodos de recessão (CHAIA, 2003).

1.2 Problema de pesquisa

Apesar da relevância do tema e da preocupação da literatura em relação à previsão de insolvência, a elaboração de trabalhos atuais se faz necessária. Isso ocorre porque as empresas contam com especificidades como setor de atuação, tamanho e forma legal (FEDOROVA; KHRUSTOVA; CHEKRYZOV, 2018). Tal fato faz com que um único modelo de previsão não seja capaz de ter um poder preditivo para qualquer empresa (FEDOROVA; KHRUSTOVA; CHEKRYZOV, 2018).

Deve-se também levar em consideração que diferentes contextos influenciam na insolvência empresarial, mas a literatura se concentrou principalmente nas análises em países desenvolvidos (OZ; SIMGA-MUGAN, 2018). É necessário, portanto, estender a literatura usando contextos diferentes dos tradicionalmente usados, uma vez que as características institucionais específicas de cada país (CUERVO-CAZURRA, 2016) promovem resultados diferentes dos até então encontrados. Sendo assim, modelos atuais devem ser elaborados e para a geração de conhecimento em relação as economias emergentes, como o Brasil.

De maneira complementar, a identificação de qual técnica fornece os melhores resultados de previsão, dentre as diversas que podem ser usadas na previsão de insolvência se faz apropriada. Conforme já mencionado, os estudos considerando dados de empresas brasileiras se concentraram principalmente na elaboração de modelos preditivos, sem enfatizar uma comparação entre diferentes técnicas, sendo necessário uma ampliação do campo de estudos nesse sentido (SANTOS, 2021). Isso se deve ao fato de que existem evidências de que diferentes técnicas de previsão fornecem diferentes níveis de assertividade (BARBOZA; KIMURA; ALTMAN, 2017).

Ademais, uma vez que a pandemia de COVID-19 foi um evento não previsto e que influencia no poder preditivo das técnicas de previsão (KITOWSKI; KOWAL-PAWUL; LICHOTA, 2022), considerar esse período também se faz relevante. Sendo assim, no que se refere especificamente à previsão de insolvência, este estudo contribui para a literatura ao comparar diferentes técnicas de previsão usando dados de empresas brasileiras e considerando o período de pandemia.

O entendimento sobre quais fatores são determinantes para o risco de insolvência também se faz relevante. Isso se deve ao fato de que, apesar da relevância de tal conhecimento, a literatura tem se concentrado principalmente na elaboração de modelos preditivos (BRYAN; FERNANDO; TRIPATHY, 2013). Sendo assim, a elaboração de estudos sobre fatores que interferem no risco de insolvência também são necessários, considerando empresas e contextos diferentes dos usualmente considerados pela literatura. Diante do exposto, este estudo visou responder ao seguinte **problema de pesquisa**: Quais são os riscos de insolvência que as empresas brasileiras não financeiras de capital aberto negociadas na Brasil, Bolsa, Balcão (B3), estão sujeitas a partir das perspectivas da previsão de insolvência e da estrutura de capital?

1.3 Objetivo geral

Para aprofundar o conhecimento sobre insolvência empresarial, a pesquisa tem como **objetivo geral** analisar o risco de insolvência, sob as perspectivas da previsão de insolvência e da estrutura de capital para as empresas brasileiras de capital aberto não financeiras negociadas na Brasil, Bolsa, Balcão (B3), bem como a assertividade de modelos de previsão para empresas que entraram em situação de insolvência durante a pandemia de COVID-19 e a influência da pandemia na relação entre estrutura de capital e risco de insolvência.

1.4 Objetivos específicos

Para alcançar o objetivo geral, serão desenvolvidos os seguintes objetivos específicos, que irão compor cada um dos artigos que serão realizados:

- (Artigo 1) Realizar uma revisão bibliométrica sobre previsão e risco de insolvência;
- (Artigo 2) Identificar entre diferentes técnicas estatísticas e de aprendizado de máquina qual fornece resultados mais robustos de previsão de insolvência considerando também o período de pandemia;
- (Artigo 3) Analisar o efeito da estrutura de capital no risco de insolvência e o impacto da pandemia no efeito da estrutura de capital no risco de insolvência.

1.5 Justificativas de pesquisa

A relevância deste estudo abrange diferentes perspectivas. Conforme mencionado, as consequências da insolvência empresarial afetam diversos agentes de toda a sociedade (SHI;

LI, 2019a). Face a isso, todos os agentes que possuem ligação com a empresa que pode tornar-se insolvente se beneficiarão com o maior entendimento sobre os fatores que possam influenciar no risco de falência.

Primeiramente, pode-se pensar em uma **relevância gerencial**, em que os benefícios para os tomadores de decisão sejam citados. A partir do maior conhecimento sobre o risco de insolvência, esses tomadores de decisão irão traçar estratégias que evitam que a empresa chegue à situação de liquidação, fato que não será benéfico para o próprio tomador de decisão nem para os demais funcionários da organização, uma vez que a manutenção dos empregos dependerá, dentre outros fatores, da continuidade da vida da empresa.

A **relevância política** também pode ser mencionada. Isso se deve ao fato de que o conhecimento sobre os fatores que influenciam o risco de insolvência é utilizado para a definição de políticas públicas que irão mitigar tal risco, e para a adequação da legislação sobre recuperação judicial, extrajudicial e falência. Assim, o estudo poderá auxiliar no esforço político para evitar a falência de empresas e minimizar suas consequências.

Quanto a **relevância social**, ao passo que a não falência de uma empresa traz benefícios para a sociedade como um todo. Uma vez que fornecedores e funcionários, por exemplo, são dependentes de uma organização, a falência de uma única empresa causa um efeito sistêmico em determinada economia (BATTISTON *et al.*, 2007). Sendo assim, identificar os fatores de influência no risco de falência faz-se relevante para toda a sociedade, uma vez que uma única falência pode impactar toda a cadeia produtiva da qual aquela empresa fazia parte, impactando também a economia local, devido ao aumento no desemprego e também no efeito que pode causar nos empregos indiretos gerados pela empresa.

Por fim, a **relevância acadêmica** do tema deve ser ressaltada. Através do estudo será possível expandir o conhecimento sobre risco de insolvência. A contribuição do estudo nesse sentido será dupla, pois serão analisadas diferentes técnicas de previsão para a identificação daquela que oferece maior robustez, bem como a análise do efeito da estrutura de capital no risco de insolvência. Uma vez que maioria dos autores esteve focada em desenvolver modelos preditivos (BRYAN; FERNANDO; TRIPATHY, 2013), a elaboração de estudos que auxiliem no entendimento de fatores que possam ser determinantes para tal risco faz-se necessária.

Ademais, a literatura será estendida analisando o risco de insolvência em uma economia emergente (Brasil). Uma vez que o desenvolvimento dos estudos se deu principalmente em economias desenvolvidas (OZ; SIMGA-MUGAN, 2018) e devido à existência de fatores específicos dos contextos institucionais (CUERVO-CAZURRA, 2016), que podem influenciar na estrutura de capital e no risco de insolvência, o desenvolvimento de análises específicas em

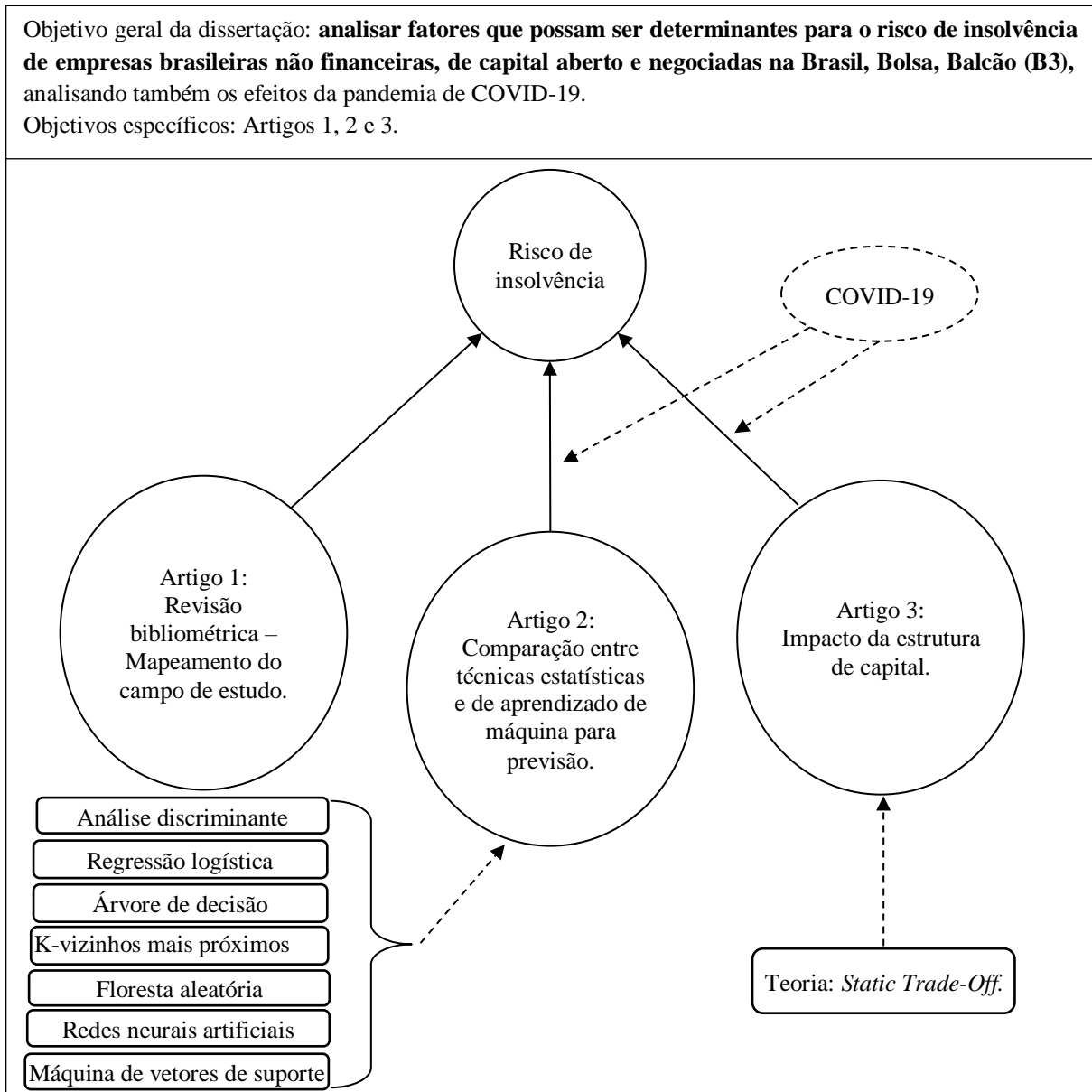
países emergentes se faz relevante. Isso se deve ao fato de que, face às diferenças institucionais entre países como os da América Latina, países Asiáticos e países desenvolvidos, diferentes resultados podem ser encontrados.

Especificamente, o Brasil foi o país emergente escolhido para a análise. Essa escolha se baseou na importância do país da América Latina, que possui o maior PIB percentual da região, correspondente a 33,25% do PIB total da região no ano de 2021 (CEPAL, 2021). A escolha também se baseou na disponibilidade de dados confiáveis para as análises do estudo.

1.6 Estrutura da dissertação

Para a construção do projeto de dissertação, o trabalho será dividido na parte inicial, sobre o tema geral do estudo, e a segunda parte, constituída pelos artigos que irão compor os objetivos específicos. A primeira parte será constituída por introdução, referencial teórico, e metodologia gerais do trabalho. Já a segunda parte possuirá introdução, referencial teórico e metodologia de cada artigo específico. Na Figura 1 encontra-se o modelo conceitual da dissertação, que relaciona os três artigos a serem desenvolvidos.

Figura 1 — Modelo conceitual da dissertação.



Fonte: Da autora (2023).

Por meio da Figura 1, percebe-se que os três artigos tiveram como tema central o risco de insolvência (ou risco de falência). Inicialmente, foi realizado um trabalho de revisão para mapear o desenvolvimento do campo de estudos. O segundo artigo objetivou a elaboração de um modelo preditivo e a comparação de técnicas de previsão. Para tentar expandir o conhecimento sobre fatores que impactam o risco de insolvência, no terceiro artigo o efeito da estrutura de capital nesse risco foi analisado. Por fim, menciona-se que o período pandêmico causado pela COVID-19 foi considerado, para verificar seus efeitos no poder preditivo das técnicas de previsão, bem como sobre a relação entre a estrutura de capital e risco de insolvência.

2 REFERENCIAL TEÓRICO

Neste tópico foi apresentado o referencial teórico geral do estudo. Para melhor compreensão do assunto se faz necessário o entendimento sobre conceitos importantes para a temática. Sendo assim, para embasar o estudo, estão explicados abaixo os conceitos de insolvência empresarial, insolvência de saldo e estrutura de capital, risco de insolvência e previsão de insolvência.

2.1 Insolvência empresarial

As dificuldades financeiras são um problema que podem levar as empresas a chegar em situações de insolvência. Lemes Júnior, Rigo e Cherobim (2010) afirmam que as dificuldades financeiras são decorrentes de fatores financeiros, contábeis, econômicos ou legais, e a existência dessas dificuldades pode comprometer o pagamento das obrigações dentro do prazo de vencimento. Corroborando o mencionado, de acordo com Mateus (2010) uma empresa insolvente é aquela que passa a ter problemas em relação à sua capacidade de pagamento (insuficiência de caixa). Ou seja, quando as empresas não conseguem gerar o fluxo de caixa suficiente para quitar suas obrigações com terceiros dentro do prazo de vencimento, elas se tornam insolventes.

Apesar dos efeitos negativos das dificuldades financeiras para as empresas, elas recorrem à justiça para sua reorganização ou liquidação. No Brasil, a Lei de falências (Lei nº 11.101) regulamentava a recuperação judicial, recuperação extrajudicial e falência. De acordo com o Diário Oficial da União (2021) a Lei Nº 14.112, de 24 de dezembro de 2020 alterou as Leis nº 11.101, de 9 de fevereiro de 2005, 10.522, de 19 de julho de 2002, e 8.929, de 22 de agosto de 1994. A alteração atualizou a legislação referente à recuperação judicial, recuperação extrajudicial e à falência.

Mesmo diante da possibilidade de recuperação extrajudicial, recuperação judicial e falência, essas categorias possuem diferenças entre si. A recuperação extrajudicial, de acordo com Lemes Júnior, Rigo e Cherobim (2010), permite que o devedor negocie com seus credores fora dos tribunais. Para isso, o devedor deverá elaborar um plano de pagamento de dívidas que deve ser aprovado pelos credores e, posteriormente, este plano deverá ser homologado no judiciário.

Na recuperação judicial, por sua vez, tem-se o desenvolvimento de um plano de recuperação, elaborado por empresas que estejam enfrentando dificuldades, sendo uma forma

de dar continuidade às atividades, evitando a falência (JUPETIPE *et al.*, 2017). Castro Júnior (2003) complementa afirmando que na recuperação judicial há a solicitação (judicial) de trégua nos pagamentos diante da proposta de um planejamento para o pagamento posterior das obrigações dentro de um período de tempo. Na recuperação judicial, segundo o autor, a empresa continua sendo comandada pelo empresário.

Por fim, a falência pode ser entendida como um processo jurídico em que o empresário não comanda mais a empresa, que passa a ser comandada pelo chamado síndico. O síndico é nomeado pelo juiz, e administra a massa falida até que o processo de liquidação termine (LEMES JÚNIOR; RIGO; CHEROBIM, 2010). Nesse caso, portanto, há o encerramento das atividades da empresa.

2.2 Insolvência de saldo e Estrutura de Capital

Além da insolvência relacionada à dificuldade de pagamento, decorrente de geração de fluxo de caixa insuficiente, existe a chamada insolvência de saldo, relacionada à estrutura de capital da empresa. Esse tipo de insolvência é caracterizado pela existência de patrimônio líquido negativo.

De acordo com Silva (2010), quando o valor do passivo total é maior que o valor do ativo total, a empresa possui insuficiência de saldo, ou seja, patrimônio líquido negativo. Isso significa que, em caso de liquidação, o valor da empresa não será suficiente para pagar suas obrigações com terceiros e remunerar os acionistas. Essa visão, de maneira similar à de Merton (1974), considera a estrutura de capital da empresa, no entanto, no caso da insolvência de saldo, não se considera o valor de mercado, mas sim o valor contábil da empresa, sendo considerada insolvente aquela que apresenta um valor total do ativo menor que o valor da dívida.

Brito, Corrar e Batistella (2007) afirmam que a estrutura de capital é um dos temas centrais na teoria das finanças corporativas, sendo amplamente discutido. Segundo os autores, as discussões têm se preocupado principalmente em definir se a estrutura de capital da empresa influencia o valor da empresa ou não. Em resumo, as empresas financiam seus ativos por meio de capital próprio ou capital de terceiros. O capital próprio diz respeito ao capital investido por sócios e acionistas, enquanto o capital de terceiros é obtido por meio do endividamento, e em ambos os casos existe o chamado custo de capital.

A literatura empírica sobre estrutura de capital possui, no entanto, visões distintas. Segundo Brito, Corrar e Batistella (2007), a teoria de estrutura de capital possui duas correntes principais, que são a corrente tradicionalista (Teoria convencional) e a corrente iniciada por Modigliani e Miller (1958). Essas correntes defendem diferentes pontos de vista sobre como a

estrutura de endividamento das empresas é definida e sobre qual a influência da estrutura de capital no valor da empresa.

De acordo com a teoria convencional, a estrutura de capital é relevante para o valor da empresa (BRITO; CORRAR; BATISTELLA, 2007; DURAND, 1959). Segundo essa corrente existe um ponto ótimo, ou estrutura de capital ótima. Sendo assim, a empresa apresentaria um ponto ótimo em que o custo de capital de terceiros e o custo do capital próprio fossem minimizados e o valor da empresa maximizado (menor custo de capital possível). Isso significa que o nível de endividamento, segundo a corrente tradicionalista é um fator que impacta no valor empresarial. Sendo assim, os tomadores de decisão devem dar atenção ao grau de endividamento das empresas, para que o valor da empresaseja maximizado (DURAND, 1959).

Modigliani e Miller (1958) introduziram a ideia de que a estrutura de capital não impactaria o valor das empresas, uma visão que contrapõe o proposto pela teoria convencional. Segundo Modigliani e Miller (1958) o grau de endividamento seria irrelevante para o valor das empresas. Para eles, o custo de capital seria o mesmo independentemente do nível de capital de terceiros, ou seja, do nível de endividamento. Sendo assim, os autores defendiam que não existe uma estrutura de capital ótima que maximiza o valor da empresa. Para eles, o valor da empresa dependeria da capacidade da empresa em gerar fluxos de caixa, bem como de seu nível de risco (BRITO; CORRAR; BATISTELLA, 2007).

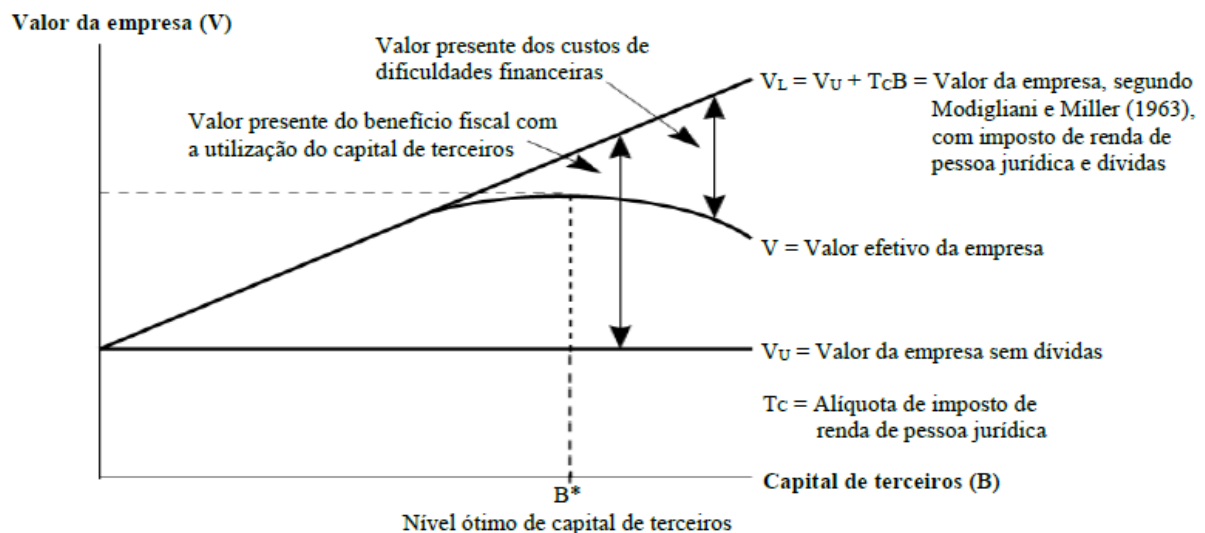
Posteriormente, os autores modificaram essa visão da irrelevância da estrutura de capital. No artigo de 1963, Modigliani e Miller assumiram que o uso de capital de terceiros proporcionaria um efeito de escudo fiscal que diminuiria o imposto de renda pago pelas organizações. Ou seja, o valor referente aos juros sobre o de capital de terceiros seria deduzido na apuração do imposto de renda, fazendo com que o valor do imposto pago fosse reduzido. Sendo assim, a ideia proposta pelos autores era a de que o uso de capital de terceiros poderia maximizar o valor das empresas. Apesar do mencionado, diferente da teoria convencional, Modigliani e Miller (1963) não consideraram os custos de falência, que posteriormente passaram a ser considerados (MYERS, 2001).

A partir da ideia de Modigliani e Miller (1963), a teoria de estrutura de capital foi sendo desenvolvida (HARON *et al.*, 2021). Trabalhos posteriores começaram a considerar o benefício fiscal e os custos de falência como relevantes para a maximização do valor da empresa (DEANGELO; MASULIS, 1980; MILLER, 1977; SCOTT, 1976). Existem, basicamente, três teorias principais sobre a estrutura de capital que se desenvolveram ao longo do tempo. Elas são a teoria do *trade-off*, a teoria do *pecking order* e a teoria do *free cash flow*. Essas teorias

ênfatizam, respectivamente, os impostos, a assimetria de informações e os custos de agência (MYERS, 2001).

A teoria em que esse estudo se baseou foi a teoria do *trade-off*, também conhecida como teoria dos custos de falência. A teoria do *trade-off* foi escolhida, pois ela considera os efeitos da estrutura de capital no valor da empresa mediante o aumento dos custos de falência, decorrentes do aumento do risco de insolvência (MYERS, 2001), e este último é considerado como o tema central da presente dissertação. De acordo com essa teoria, o benefício fiscal proveniente do uso de capital de terceiros aumenta o valor da empresa até certo ponto, que é onde os custos decorrentes do risco de insolvência se tornam superiores a esse benefício devido ao aumento do nível de endividamento, conforme representado pela Figura 2.

Figura 2 — Teoria do *trade-off*.



Fonte: Adaptado de Ross, Westerfield, Jaffe (2007, p. 432).

Por meio da Figura 2, é possível verificar que, de acordo com a teoria do *trade-off* o valor da empresa é influenciado pela estrutura de capital (especificamente, pelo nível de endividamento). Mesmo considerando o benefício fiscal decorrente do uso de capital de terceiros (MODIGLIANI; MILLER, 1963), é necessário considerar os custos de falência. Sendo assim, de acordo com essa teoria existe um ponto ótimo da estrutura de capital, que maximiza o valor da empresa, principalmente devido ao benefício fiscal que diminui o imposto de renda pago.

A partir do ponto-ótimo, no entanto, os custos de falência associados ao aumento do risco de insolvência, decorrente do aumento do endividamento, superam o benefício fiscal do

uso de capital de terceiros. Ou seja, a partir do ponto ótimo, o aumento no nível de endividamento causa um impacto negativo no valor da empresa.

Apesar da existência do benefício fiscal, de acordo com a teoria do *trade off*, é importante destacar que este só ocorre mediante a incidência de juros sobre o capital de terceiros. Uma vez que, existem diferentes tipos de passivo, ressalta-se que em alguns casos não há a incidência explícita de juros (passivo não oneroso), sendo assim, é importante considerar essas diferenças, uma vez que o passivo oneroso é o que se caracteriza pela incidência de despesas de juros (MACHADO; MEDEIROS; EID JÚNIOR, 2010). Logo, espera-se que o passivo oneroso seja o responsável pelo benefício fiscal gerado pelo endividamento.

2.3 Risco de insolvência (ou risco de falência)

A definição do que é o risco é considerado um aspecto importante para sua avaliação (HORVÁTHOVÁ; MOKRIŠOVÁ, 2018). De acordo com Horváthová e Mokrišová (2018), o risco é entendido como a ocorrência de um evento inesperado, sendo tratado como algo indefinido e instável, que causa uma perturbação no comportamento de determinado fenômeno. Sendo assim, o risco é tratado como uma incerteza, caracterizada pela probabilidade de desvio dos valores esperados, sendo este desvio positivo ou negativo. No entanto, faz-se necessário destacar que na prática o risco é considerado como um desvio negativo (HORVÁTHOVÁ; MOKRIŠOVÁ, 2018).

O risco de falência (ou risco de insolvência) é entendido como a probabilidade que uma empresa entre em situação de insolvência, fato que culmina em sua falência. Ou seja, o risco de falência é entendido como a probabilidade de que uma empresa não seja capaz de quitar suas obrigações conforme seu vencimento (MATEUS, 2010). Face às possíveis consequências do risco de insolvência, esse campo despertou a atenção de estudiosos ao longo das últimas décadas.

De acordo com Bryan, Fernando e Tripathy (2013), a falência é um evento importante no ambiente empresarial moderno, uma vez que a ela pode desencadear um efeito negativo em seus *stakeholders*. Segundo os mesmos autores, a falência ocorre quando uma empresa solicita judicialmente um período para que possa se reorganizar financeiramente ou encerrar suas atividades.

Apesar da especificidade do evento de falência, as dificuldades financeiras geralmente têm início bem antes desse evento. Diante desse fato, estudiosos elaboraram modelos preditivos

capazes de antecipar situações de insolvência ou falência (BRYAN; FERNANDO; TRIPATHY, 2013). Nesse sentido, a previsão de insolvência (ou previsão de falência) é entendida como a antecipação de tal evento.

Por meio da previsão, as empresas tomam estratégias corretivas para evitar a falência, e a previsão pode auxiliar nas decisões de concessão de crédito e investimentos. Face à relevância dos estudos de previsão, diversos estudos sobre a previsão falência foram desenvolvidos, como os de Beaver (1966), Altman (1968) e Ohlson (1980). Entre os estudos brasileiros, são citados os trabalhos de Kanitz (1974), Elizabetsky (1976), Altman, Baidya e Dias (1979), Silva (1982), Sanvicente e Minardi (1998), Stüpp (2015) e Prado *et al.* (2020), que usaram diferentes técnicas de previsão, conforme detalhado no Quadro 1.

Quadro 1 — Resumo de estudos sobre previsão de insolvência (Continua).

Autores	Técnica	Nível de assertividade
Altman (1968)	Análise discriminante	95%.
Kanitz (1974)	Regressão múltipla e análise discriminante	Media o nível de cada empresa individualmente (com a aplicação menor rejeição de financiamentos e empréstimos).
Elizabetsky (1976)	Regressão multilinear	74,07% das empresas boas corretamente classificadas e 62,96% das empresas ruins.
Matias (1978)	Análise discriminante	Das 50 empresas solventes, 44 foram classificadas corretamente (88%). Das empresas insolventes, 45 foram classificadas de forma correta (90%).
Altman, Baidya e Dias (1979)	Análise discriminante	78% (três anos antes da situação-problema); 88% (um ano de antecedência).
Ohlson (1980)	Regressão logística	85,1%.
Silva (1982)	Análise discriminante	90% - Empresas solventes e 83% - empresas insolventes.
Sanvicente e Minardi (1998)	Análise discriminante	81,8%
Akiama (2008)	Regressão logística	72,7%
Stüpp (2015)	Análise discriminante e regressão logística	Análise discriminante: acerto de 76,1%; Regressão logística: acerto de 76,2%.
Prado <i>et al.</i> (2020)	Redes Neurais artificiais	Modelo MT: 96,9% de acerto global; Modelo MT&D: 100% de acerto.

Quadro 1 — Resumo de estudos sobre previsão de insolvência (conclusão).

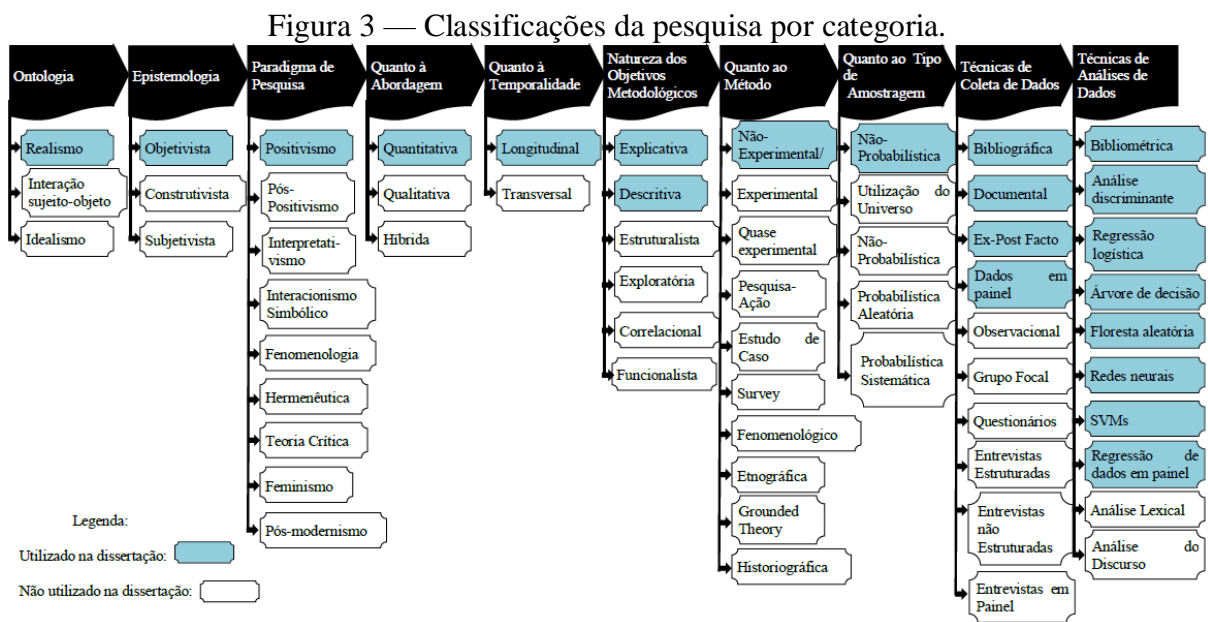
Pinto (2021)	Naive Bayes SVM Linear SVM Polynomial SVM Radial Regressão logística Redes neurais Floresta aleatória XGBoost	Naive Bayes: 94,96% SVM Linear: 95,28% SVM Polynomial: 95,09% SVM Radial: 96,51% Regressão logística: 96,55% Redes neurais: 96,38% Floresta aleatória: 97,8% XGBoost: 98,38%
Santos (2021)	Floresta aleatória, <i>naive bayes</i> , <i>logit</i> , K-NN, SVM (linear, polinomial e de base radial), <i>bagging</i> e <i>boosting</i>	Floresta aleatória: 95,72% <i>naive bayes</i> : 91,01% <i>logit</i> : 92,51% K-NN: 94,01% SVM (linear, polinomial e de base radial): 92,8%, 93,72% e 95,22% <i>Bagging</i> : 94,44% <i>Boosting</i> : 94,86%

Fonte: Da autora (2023).

Como pode ser percebido no Quadro 1, os estudos brasileiros na maioria dos casos usaram apenas uma técnica de previsão, com ênfase na definição de um modelo preditivo e não de comparação de técnicas. Apesar de nos últimos anos alguns autores terem usado mais de uma técnica (STÜPP, 2015; PINTO, 2021; SANTOS, 2021), ainda percebe-se uma escassez de trabalhos brasileiros com que enfatizam essa comparação, sendo necessário o desenvolvimento de novos estudos, considerando inclusive o período de pandemia.

3 METODOLOGIA GERAL

Neste tópico a metodologia geral do estudo será delineada. Os estudos possuem diversos tipos de classificações, que, conforme Prado (2019) são categorizadas quanto à ontologia, epistemologia, paradigma de pesquisa, abordagem, temporalidade, natureza dos objetivos metodológicos, método, tipo de amostragem, técnicas de coleta de dados e técnicas de análise de dados. Na Figura 3 encontram-se as classificações do estudo conforme cada uma das categorias mencionadas.



Fonte: Adaptado de Prado (2019, p. 56).

De acordo com Saccol (2009) a ontologia pode ser entendida como a compreensão do pesquisador sobre como as coisas são, ou seja, está ligada à questão do “ser”. De maneira complementar Souza, Costa e Pereira (2015, p. 732) afirmam que a perspectiva ontológica “discute e problematiza a natureza da realidade”. De acordo com Saccol (2009), a posição ontológica assumida pelos indivíduos é responsável pela definição da forma como o mundo e os fenômenos sob análise são vistos, existindo duas visões opostas, que são a realista e a idealista. Neste caso, a classificação do estudo quanto à ontologia é realista. O realismo pressupõe, segundo Saccol (2009, p. 253) a existência de “um mundo lá fora” que é “independente das percepções e construções mentais” que possam ser criadas sobre ele.

Ao considerar a epistemologia, o estudo se classifica como objetivista. A epistemologia pode ser considerada como a forma por meio da qual o conhecimento é gerado no entendimento

dos indivíduos, estando fortemente ligada aos pressupostos ontológicos, de “como as coisas são” (SACCOL, 2009, p. 253). Existem, de acordo com Saccol (2009), três grandes de linhas de pensamento nessa categoria, que são subjetivismo, objetivismo e construtivismo. Este estudo pode ser classificado como objetivista que, entendida como uma perspectiva em que se “pressupõe que os significados sobre todos os objetos e entidades existem independentemente de operações mentais do ser humano”, em que tudo que existe possui um significado objetivo (SACCOL, 2009, p. 254).

Outra categoria destacada foi a de paradigma de pesquisa. De acordo com Saccol (2009), os paradigmas de pesquisa incluem a existência de visões de mundo diferentes. Ainda de acordo com a autora, existem diferentes paradigmas de pesquisa, que, em resumo, partem de instâncias filosóficas e definem a metodologia de pesquisa. Este estudo pode ser classificado, nesta categoria, como positivista. O positivismo pode ser considerado como fortemente ligado a pesquisas quantitativas, se fundamentando na ontologia realista, tendo como base, portanto, a crença em verdades objetivas, que não dependem da percepção dos seres humanos (SACCOL, 2009).

Ao tratar da abordagem, o estudo pode ser classificado como quantitativo. De acordo com Appolinário (2016, p. 23), na pesquisa quantitativa há a “mensuração de variáveis pré-determinadas, buscando verificar e explicar sua influência sobre outras variáveis”, tendo como ênfase informações matematizáveis, por meio das quais generalizações sobre o fenômeno estudado podem ser realizadas. Sendo assim, os três artigos contam com abordagens quantitativas, por meio de indicadores bibliométricos sobre a área temática ou indicadores econômico-financeiros extraídos a partir das demonstrações financeiras das empresas analisadas.

No que diz respeito à temporalidade, o estudo se classifica como longitudinal. De acordo com Prado (2019), os estudos longitudinais são caracterizados por analisar fenômenos ao longo do tempo. Uma vez que na análise bibliométrica tem-se o mapeamento do campo de pesquisa desde o início das publicações até o período atual, e nos modelos de previsão e no estudo da influência da estrutura de capital no risco de falência o período considerado será de 2010 a 2021, justifica-se a classificação do estudo como longitudinal.

Em relação à natureza dos objetivos metodológicos, o estudo se classifica como explicativo e descritivo. Essa classificação se dá devido ao fato de que o estudo irá descrever e explicar o fenômeno analisado, que é o risco de falência (MATIAS-PEREIRA, 2019). Na pesquisa explicativa, há a intenção de elucidação de fatores que podem influenciar um

fenômeno, tendo como base a pesquisa descritiva, que descreve as características do fenômeno analisado, estabelecendo relações entre as variáveis (VERGARA, 2016).

Quanto ao método, pode ser classificado como não-experimental. Essa classificação se deve ao fato de que as variáveis não são controladas, para que seus efeitos sobre o objeto de estudo sejam identificados (APPOLINÁRIO, 2016).

Ao tratar do tipo de amostragem, o presente estudo pode ser classificado como não-probabilístico e intencional, uma vez que os indivíduos analisados serão escolhidos intencionalmente (LAKATOS, 2019). Para o estudo de bibliometria (primeiro artigo) será realizado um levantamento de outros estudos que fazem parte da temática, escolhidos através da *string* de buscas, para que o mapeamento da produção científica seja feita a partir de tais estudos.

Para os modelos de previsão, serão usadas empresas não financeiras de capital aberto e com ações negociadas na B3. A amostra do segundo artigo será composta por empresas insolventes caracterizadas por apresentar situação de recuperação judicial, recuperação extrajudicial, falência ou patrimônio líquido negativo, bem como uma amostra equiparada de empresas solventes, de preferência que sejam do mesmo setor e com o ativo total próximo ao da respectiva empresa insolvente.

Por fim, para o estudo do efeito da estrutura de capital no risco de insolvência (terceiro artigo), serão consideradas empresas não financeiras de capital aberto ativas com ações negociadas na B3. A análise será feita por meio da análise de dados em painel, com todas as empresas que apresentarem as informações suficientes para a análise.

No que se refere às técnicas de coleta de dados, será usada a pesquisa documental, caracterizada como *ex-post facto*. A pesquisa documental pode ser entendida como a pesquisa em que são realizadas consultas a documentos (CERVO; BERVIAN; SILVA, 2007), como as demonstrações financeiras, exemplo deste estudo. A classificação *ex-post facto* se deve ao fato de que as informações que serão usadas no estudo não são manipuláveis, uma vez que já ocorreram (VERGARA, 2016).

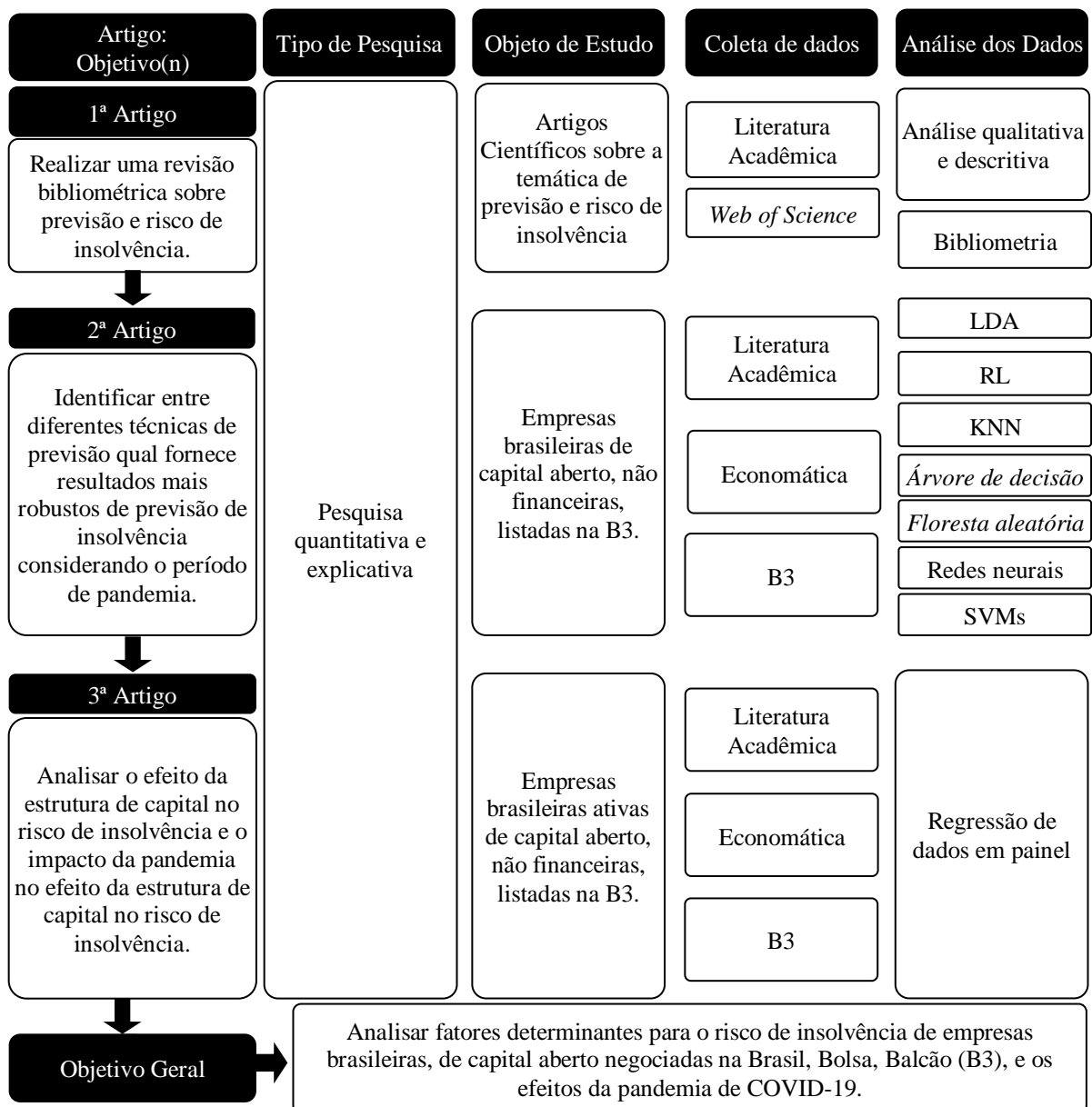
Por fim, quanto às técnicas de análise de dados, são citadas a bibliometria (primeiro artigo); análise discriminante, regressão logística, árvores de decisão, floresta aleatória, redes neurais artificiais e máquinas de vetores de suporte (segundo artigo); e regressão de dados em painel, (terceiro artigo).

A bibliometria é entendida como uma técnica quantitativa de análise, usada para o mapeamento do campo de estudos (VANTI, 2002). A análise discriminante e a regressão logística são técnicas estatísticas usadas para a previsão, enquanto árvores de decisão, floresta

aleatória, redes neurais artificiais e máquinas de vetores de suporte são técnicas de aprendizado de máquina (IZBICK; SANTOS, 2020). Por fim, a regressão de dados em painel é uma técnica que usa dados de vários indivíduos em períodos de tempo diferentes (GUJARATI; PORTER, 2011).

Na Figura 4 está uma síntese dos procedimentos metodológicos do estudo. Como a dissertação foi organizada em artigos, em cada um dos artigos a metodologia foi explicada com maiores detalhes. Sendo assim, a Figura 4 mostra apenas resumidamente as etapas de cada artigo, no que se refere à metodologia empregada em cada um deles.

Figura 4 — Resumo da proposta de metodologia para a dissertação.



Fonte: Adaptado de Antonialli (2018, p. 54).

Uma vez que no artigo 2 foram usadas diferentes técnicas e seus resultados foram comparados, fez-se necessário o uso de um parâmetro para verificar se tais técnicas são confiáveis. Para isso, a curva ROC (*Receiver Operating Characteristic Curves*) de cada modelo foi analisada. De acordo com Sicsú (2010), essa curva é baseada nas dimensões de sensibilidade e especificidade, e a principal medida a ser observada é a área sob a curva (AUC: *Area Under the Curve*). De acordo Soares e Rebouças (2014), essa área (AUC) varia de zero a um, sendo que quanto mais próxima de um, maior é o ajuste do modelo, sendo, portanto, melhor que os valores da AUC sejam mais próximos de um.

4 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo analisar o risco de insolvência de empresas brasileiras não financeiras listadas na B3, sob as perspectivas da previsão de insolvência e estrutura de capital, analisando a assertividade de modelos de previsão durante a pandemia de COVID-19 e os efeitos da COVID-19 na influência da estrutura de capital sobre o risco de insolvência. Para alcançar o objetivo geral, os objetivos específicos foram desenvolvidos em forma de artigos, sendo eles: Realizar uma revisão bibliométrica sobre previsão e risco de insolvência (artigo 1); identificar entre diferentes técnicas estatísticas e de aprendizado de máquina qual fornece resultados mais assertivos para o período de pandemia (artigo 2); analisar o efeito da estrutura de capital no risco de insolvência e o impacto da pandemia como moderadora nessa relação (artigo 3).

Por meio do artigo 1, intitulado “*Risco e previsão de falência: Uma revisão bibliométrica*”, foi possível realizar o mapeamento do campo de estudos, mostrando que o campo de estudos se desenvolveu lentamente principalmente até os anos 2000, com uma tendência de crescimento a partir de 2008, provavelmente associada à crise do subprime. Posteriormente, houve aumentos consideráveis a partir do ano de 2016, com auge de publicações em 2022, que provavelmente ocorreu devido à crise decorrente da pandemia de COVID-19.

O estudo mais relevante foi o de Altman (1968). Ao se tratar do volume de publicações, os autores principais foram Sun e Li. O principal país do campo de estudos, no que se refere às citações, foi os Estados Unidos, mas o país com maior volume de publicações foi a China. O periódico mais relevante foi o *Expert Systems With Application*, e ao analisar as palavras-chave do campo de estudo, as principais foram *classification, financial ratios, risk, neural-networks, models, discriminant-analysis, model, performance e prediction*. Os tópicos de tendência foram computacionais como *big data, machine learning e deep learning*, o que mostra que as técnicas de aprendizado de máquina, usadas no artigo 2, são uma tendência no que se refere aos modelos de previsão de insolvência.

No artigo 2, nomeado “*Previsão de insolvência: Uma análise comparativa de diferentes técnicas de previsão para o momento de pandemia de COVID-19*”, técnicas estatísticas (análise discriminante, regressão logística) e de aprendizado de máquina das três classificações principais: Caixa branca (k-vizinhos mais próximos e árvore de decisão), caixa cinza (floresta aleatória) e caixa preta (RNAs e SVMs) foram analisadas, considerando o período de pandemia como amostra de teste para avaliar o desempenho dos modelos de previsão. Como resultado, a

técnica com melhor acurácia foi a de árvore de decisão (DT), seguida pela de floresta aleatória (RF), regressão logística e análise discriminante. Por último estiveram as técnicas de rede neural artificial (RNA), k-vizinhos mais próximos (KNN) e máquina de vetores de suporte (SVM).

Esses resultados mostraram que as técnicas de caixa branca, caixa cinza e estatísticas tiveram melhores acurácias se comparadas às técnicas de caixa preta (RNA e SVM). Sendo assim, considerando as empresas que entraram em situação de insolvência no período de pandemia como amostra de teste, os resultados apontaram a superioridade de uma técnica de caixa branca (DT), seguida de uma de caixa cinza (RF) e das técnicas estatísticas tradicionais (LDA e LR) sobre as técnicas de caixa preta. Esses resultados mostraram que, diferentemente do que geralmente é esperado (OBERMANN; WAACK, 2015), as técnicas de caixa preta (RNA e SVM) apresentaram os piores desempenhos de acurácia.

Devido à grande atenção da literatura aos modelos de previsão, a elaboração de estudos para identificar os fatores de influência no risco de insolvência também são necessários. Sendo assim, o artigo 3, intitulado “*Efeitos da estrutura de capital no risco de insolvência*” teve como objetivo analisar a influência da estrutura de capital no risco de insolvência de empresas brasileiras não financeiras de capital aberto negociadas na B3, verificando também o efeito moderador da pandemia de COVID-19 nessa relação. Por meio de regressão multinível de dados em painel, foi possível identificar que algumas das variáveis de estrutura de capital apresentaram resultados significativos e positivos, confirmando as primeiras hipóteses propostas (especificamente, as variáveis, endividamento de curto prazo (END_C_LP) e endividamento de longo prazo (END_C_LP)), resultados que confirmaram as hipóteses 1b e 1c.

O resultado da variável dívida líquida de curto prazo (DIV_LIQ_CP), por sua vez foi significativo porém negativo, diferente do esperado. Logo a hipótese 1a do estudo não foi confirmada, indicando que a dívida líquida de curto prazo influenciou negativamente no risco de insolvência das empresas analisadas.

Ao analisar a variável endividamento financeiro a valor contábil (END_C_DF) e o termo quadrático dessa variável (END_C_DF_2), foi possível perceber que elas apresentaram resultados significativos, e com coeficientes negativo e positivo, respectivamente. Esses resultados confirmaram a hipótese 1d do estudo, que, em que se esperava uma relação quadrática em forma de U entre o endividamento financeiro e o risco de insolvência. Esse resultado era esperado devido ao benefício fiscal advindo do passivo oneroso, que compensa, até certo ponto ($END_C_DF = 0,4679$), os custos de falência associados ao endividamento.

Ao analisar o efeito moderador da pandemia, no entanto, os resultados se mostraram significativos em alguns casos, mas não conforme o esperado por todas as hipóteses. Especificamente, apenas a hipótese de que a pandemia de COVID-19 fortalece a influência positiva do endividamento de longo prazo no risco de insolvência foi confirmada. Sendo assim, as demais hipóteses que analisavam os efeitos da pandemia de COVID-19 na relação entre dívida líquida de curto prazo, endividamento de curto prazo e endividamento financeiro no risco de insolvência não foram confirmadas.

As limitações dos estudos estão associadas às limitações individuais de cada um dos artigos desenvolvidos. Entre elas, mencionam-se o uso apenas da técnica de bibliometria para a revisão do campo de estudos (artigo 1), foram considerados dados apenas de empresas brasileiras de capital aberto com ações negociadas na B3 (Artigos 2 e 3), uso apenas de indicadores econômico-financeiros para a elaboração dos modelos preditivos (artigo 2) e da análise da influência da estrutura de capital no risco de insolvência (artigo 3).

Sendo assim, para estudos futuros sugere-se que estudos de revisão sejam realizadas utilizando a análise de conteúdo dos trabalhos, ou associando as técnicas de bibliometria e análise de conteúdo para o aprofundamento do campo de estudos. Sugere-se também que estudos futuros considerem dados de outros países (emergentes ou desenvolvidos) para a elaboração dos modelos de previsão analisando sua assertividade no momento de pandemia, e para a análise da influência da estrutura de capital no risco de insolvência e do efeito moderador da pandemia de COVID-19. Por fim, sugere-se que estudos futuros para a elaboração dos modelos de previsão e para a análise da influência da estrutura de capital no risco de insolvência considerem variáveis macroeconômicas, uma vez que os fatores do ambiente institucional se diferenciam nos diferentes contextos existentes entre os diferentes países.

REFERÊNCIAS

- ALTMAN, E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, n. 4, p. 589–609, 1968.
- ALTMAN, E.; BAIDYA, T.; DIAS, L. Previsão de problemas financeiros em empresas. **Revista de Administração de Empresas**, v. 19, n. 1, p. 17–28, 1979.
- APPOLINÁRIO, F. **Metodologia da Científica**. 2ª ed. [s.l.] Cengage Learning, 2016.
- BARBOZA, F.; KIMURA, H.; ALTMAN, E. Machine learning models and bankruptcy prediction. **Expert Systems With Applications**, v. 83, p. 405–417, 2017.
- BATTISTON, S. *et al.* Credit chains and bankruptcy propagation in production networks. **Journal of Economic Dynamics and Control**, v. 31, n. 6, p. 2061–2084, 2007.
- BEAVER, W. H. Financial Ratios as Predictors of Failure. **Journal of Accounting Research**, v. 4, p. 71–111, 1966.
- BRITO, G. A. S.; CORRAR, L. J.; BATISTELLA, F. D. Fatores determinantes da estrutura de capital das maiores empresas que atuam no Brasil. **Revista Contabilidade & Finanças**, v. 18, n. 43, p. 9–19, 2007.
- BRYAN, D.; FERNANDO, G.; TRIPATHY, A. Bankruptcy risk, productivity and firm strategy. **Review of Accounting and Finance**, v. 12, n. 4, p. 309–326, 2013.
- CASTRO JÚNIOR, F. H. F. **Previsão de Insolvência de Empresas Brasileiras Usando Análise Discriminante, Regressão Logística e Redes Neurais**. [s.l.] Universidade de São Paulo, 2003.
- CEPAL. **Anuario Estadístico de América Latina y el Caribe**. Disponível em: <<https://statistics.cepal.org/yearbook/2021/statistics.html?lang=es&theme=economicas>>. Acesso em: 13 set. 2022.
- CERVO, A. L.; BERVIAN, P. A.; SILVA, R. **Metodologia Científica**. 6ª ed. São Paulo: Pearson Prentice Hall, 2007.
- CHAIA, A. J. **Modelos de gestão do risco de crédito e sua aplicabilidade ao mercado brasileiro**. 2003. Dissertação (Mestrado em Administração) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2003.
- CUERVO-CAZURRA, A. Multilatinas as sources of new research insights: The learning and escape drivers of international expansion. **Journal of Business Research**, v. 69, n. 6, p. 1963–1972, 2016.
- DEANGELO, H.; MASULIS, R. W. Optimal capital structure under corporate and personal taxation. **Journal of Financial Economics**, v. 8, n. 1, 1980.
- DIÁRIO OFICIAL DA UNIÃO. **Lei Nº 14.112, de 24 De Dezembro de 2020**. Disponível em: <<https://www.in.gov.br/en/web/dou/-/lei-n-14.112-de-24-de-dezembro-de-2020->

310838289>. Acesso em: 2 de maio de 2022.

DURAND, D. **Costs of Debt and Equity Funds for Business: Trends and Problems of Measurement**. Conference on Research in Business Finance. **Anais...NBER**, 1952

DURAND, D. The Cost of Capital, Corporation Finance, and the Theory of Investment: Comment. **American Economic Review.**, v. 49, n. 4, p. 639–655, 1959.

ELIZABETSKY, R. **Um modelo matemático para decisão no banco comercial**. [s.l.] Universidade de São Paulo, 1976.

FEDOROVA, E. A.; KHRUSTOVA, L. E.; CHEKRYZOV, D. V. Sectoral specificities of application of Bankruptcy Forecasting models. **Strategic Decisions and Risk Management**, v. 106, n. 1, p. 52–59, 2018.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5ª ed. Brasil: McGraw Hill Brasi, 2011.

HORVÁTHOVÁ, J.; MOKRIŠOVÁ, M. Risk of bankruptcy, its determinants and models. **Risks**, v. 6, n. 4, 2018.

IZBICK, R.; SANTOS, T. **Aprendizado de máquina: uma abordagem estatística**. 1. ed. São Carlos: [s.n.].

JUPETIPE, F. K. N. *et al.* Custos de falência no Brasil comparativamente aos estudos norte-americanos. **Revista Direito FGV**, São Paulo, v.13. n. 1, p. 20-48, Jan. 2017.

KANITZ, S. Como Prever Falências de Empresas. **Revista Negócios em Exame**, p. 95–102, 1974.

KITOWSKI, J.; KOWAL-PAWUL, A.; LICHOTA, W. Identifying Symptoms of Bankruptcy Risk Based on Bankruptcy Prediction Models—A Case Study of Poland. **Sustainability**, v. 14, n. 3, 2022.

LAKATOS, M. D. A. M. E. M. **Fundamentos de metodologia científica**. 8ª ed. São Paulo: Atlas, 2019.

LEMES JÚNIOR, A.; RIGO, C.; CHEROBIM, A. **Administração Financeira: Princípios, Fundamentos e Práticas Brasileiras**. 3. ed. Rio de Janeiro: Elsevier, 2010.

MATEUS, R. S. **Análise de insolvência empresarial: uma abordagem financeira fundamentalista com aplicação do método estatístico multivariado e da técnica discriminante**. Dissertação (Mestrado em Economia) – Universidade Federal de Sergipe, São Cristóvão, 2010.

MATIAS-PEREIRA, J. **Manual de metodologia da pesquisa científica**. 4ª ed. São Paulo: Atlas, 2019.

MERTON, R. C. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. **The Journal of Finance**, v. 29, n. 2, p. 449–470, 3 jun. 1974.

- MILLER, M. H. Debt and taxes. **Journal of Finance**, v. 32, n. 2, 1977.
- MODIGLIANI, F., MILLER, M. The cost of capital, corporation finance and the theory of investment. **The American Economic Review**, p. 261–297, 1958.
- MODIGLIANI, F.; MILLER, M. H. Corporate income taxes and the cost of capital: a correction. **The American Economic Review**, v. 53, n. 3, p. 433–443, 1963.
- MYERS, S. The capital structure puzzle. **Journal of Finance**, v. 39, n. 3, p. 575–592, 1984.
- MYERS, S. C. Capital structure. **The Journal of Economic Perspectives**, v. 15, n. 2, p. 81–102, 2001.
- OBERMANN, L.; WAACK, S. Demonstrating non-inferiority of easy interpretable methods for insolvency prediction. **Expert Systems With Applications**, v. 42, n. 23, p. 9117–9128, 2015.
- OHLSON, J. A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. **Journal of Accounting Research**, v. 18, n. 1, p. 109–131, 1980.
- OZ, I. O.; SIMGA-MUGAN, C. Bankruptcy prediction models’ generalizability: Evidence from emerging market economies. **Advances in Accounting**, v. 41, n. September 2017, p. 114–125, 2018.
- PINTO, A. C. O poder preditivo dos modelos com aprendizado de máquina é superior aos modelos tradicionais para análise do risco de crédito? **Revista Debates em Economia Aplicada – REDEA**, v. 1, n. 1, 2021.
- PRADO, J. W. **Determinantes e implicações da estrutura de capital, da estrutura de propriedade e da governança corporativa: um modelo multiteórico de análise**. 2019. Tese (Doutorado em Administração) – Universidade Federal de Lavras, Lavras, 2019.
- PRADO, J. *et al.* Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). **Scientometrics, Budapeste**, v. 106, p. 1007–1029, 2016.
- PRADO, J. *et al.* Previsão da insolvência empresarial utilizando redes neurais artificiais. **Revista Gestão e Desenvolvimento, Novo Hamburgo**, v. 17, n. 2, p. 136–162, 2020.
- ROSS, S. A.; WESTERFIELD, R. W.; JAFFE, J. F. **Administração Financeira**. 2. ed. São Paulo: Atlas, 2007.
- SACCOL, A. Z. Um retorno ao básico: compreendendo os paradigmas de pesquisa e sua aplicação na pesquisa em administração. **Revista de Administração da UFSM**, v. 2, n. 2, p. 250–269, 2009.
- SANTOS, A. S. **Previsão de insolvência corporativa: uma análise de empresas brasileiras de capital aberto por meio de aprendizado de máquina**. [s.l.] Universidade Federal da Paraíba, 2021.

- SANVICENTE, A. Z.; MINARDI, A. M. A. F. Identificação de indicadores contábeis significativos para previsão de concordata de empresas. **Instituto Brasileiro de Mercado de Capitais, WorkingPaper**, São Paulo, p. 1-12, 1998.
- SCOTT, J. A Theory of optimal capital structure. **The Bell Journal of Economics**, v. 7, 1976.
- SHI, Y.; LI, X. An overview of bankruptcy prediction models for corporate firms: A systematic literature review. **Intangible Capital**, v. 15, n. 2, SI, p. 114–127, 2019a.
- SHI, Y.; LI, X. A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. **Heliyon**, v. 5, n. 12, 2019b.
- SICSÚ, A. **Credit scoring: desenvolvimento, implantação, acompanhamento**. São Paulo: Blucher, 2010.
- SILVA, A. I. S. **Um Modelo de Previsão de Insolvência Financeira: Aplicação ao Sector da Indústria Transformadora**. Dissertação (Mestrado em Economia) – Universidade da Beira Interior, Covilhã, 2010.
- SILVA, J. (1982). **Modelos para classificação de empresas com vistas a concessão de crédito**. Dissertação (Mestrado em Administração) - Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas. São Paulo, 1982.
- SILVA, P. . *et al.* Teoria da Sinalização e a Recuperação judicial: um Estudo nas Empresas de Capital Aberto Listadas na BM&FBOVESPA. **Desenvolvimento em Questão**, v. 16, n. 42, 2018.
- SJÖBERG, J. *et al.* Nonlinear Black-Box Modeling in System Identification: a Unified Overview. **Automatica**, v. 31, n. 12, p. 1691–1724, 1995.
- SOARES, R.; REBOUÇAS, S. Avaliação do desempenho de técnicas de classificação aplicadas à previsão de insolvência de empresas de capital aberto brasileiras. **Revista ADM. Made**, v. 18, n. 3, p. 40–61, 2014.
- SOUZA, E. M.; COSTA, A. S. M.; PEREIRA, S. J. N. A Organização (in)corporada: ontologia organizacional, poder e corpo em evidência. **Cadernos EBAPE.BR**, v. 13, n. 4, p. 727–742, 2015.
- STÜPP, D. R. **Previsão de insolvência a partir de indicadores contábeis: evidências de empresas listadas na BM&FBOVESPA nos anos 2004-2013**. Dissertação (Mestrado em Contabilidade) - Universidade Federal de Santa Catarina, Florianópolis, 2015.
- TAO, Q. *et al.* Does operating leverage increase firm’s profitability and bankruptcy risk? Evidence from China’s entry into WTO. **International Journal of Finance and Economics**, n. November, p. 1–17, 2020.
- THE WORLD BANK. **A Economia Mundial Deve Expandir 4% em 2021; Implementação da Vacina e Investimentos são Essenciais para Sustentar a Recuperação**.

05 Jan. 2021. Disponível em: <<https://www.worldbank.org/pt/news/press-release/2021/01/05/global-economy-to-expand-by-4-percent-in-2021-vaccine-deployment-and-investment-key-to-sustaining-the-recovery>>. Acesso em: 30 mar. 2022.

THE WORLD BANK. **O crescimento global deve desacelerar até 2023, aumentando o risco de um declínio rápido e indesejado nas economias em desenvolvimento**. 11 Jan. 2022. Disponível em: <<https://www.worldbank.org/pt/news/press-release/2022/01/11/global-recovery-economics-debt-commodity-inequality>>. Acesso em: 30 mar. 2022.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, v. 31, n. 2, p. 152–162, 2002.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 16^a ed. São Paulo: Atlas, 2016.

SEGUNDA PARTE – ARTIGOS

A partir deste tópico, foram desenvolvidos os objetivos específicos da dissertação, em forma de artigos. Sendo assim, foram desenvolvidos os seguintes artigos:

- i. Risco e previsão de insolvência: Uma revisão bibliométrica (Artigo 1).
- ii. Risco de insolvência: Uma análise comparativa de diferentes técnicas de previsão para o momento de pandemia (Artigo 2).
- iii. Efeitos da estrutura de capital no risco de insolvência (Artigo 3).

ARTIGO 1 – RISCO E PREVISÃO DE INSOLVÊNCIA: UMA REVISÃO BIBLIOMÉTRICA

RESUMO

O objetivo desse artigo foi realizar uma revisão bibliométrica sobre o tema de risco e previsão de insolvência. Para isso, foram desenvolvidos os seguintes passos: realizar uma análise sistemática sobre o tema; identificar tendências temporais de publicações, principais autores, principais países de origem, principais periódicos, principais palavras-chave, entre outros. Para o desenvolvimento do estudo foi usada a técnica de bibliometria. Especificamente, para a junção das bases de dados foi utilizado o *software* R e para as análises bibliométricas utilizaram-se o M.S. Excel e o *software* Bibliometrix. Como resultado, o campo de estudos se desenvolveu lentamente até os anos 2000, com uma tendência de crescimento a partir de 2008, possivelmente associado à crise do *subprime*. Posteriormente, houve aumentos a partir de 2016, com auge de publicações em 2022, possivelmente associado à crise causada pela COVID-19. O estudo mais relevante foi o de Altman (1968). O principal país para o campo de estudos em volume de citações foi Estados Unidos, mas a China apresentou maior volume de publicações. O periódico mais relevante foi o *Expert Systems With Application*, e ao analisar as palavras-chave, as principais foram *classification, financial ratios, risk, neural-networks, models, discriminant-analysis, model, performance e prediction*. Os tópicos de tendência foram computacionais como *big data, machine learning e deep learning*. Diante disso, a partir do estudo, foi possível realizar um importante levantamento sobre os trabalhos envolvendo a temática do risco e previsão de insolvência, utilizando um amplo conjunto de dados (*Scopus e Web of Science*).

Palavras-chave: Previsão de falência. Risco de falência. Bibliometria.

ABSTRACT

The aim of this article was to carry out a bibliometric review on the subject of risk and prediction of insolvency. For this, the following steps were developed: perform a systematic analysis on the subject; identify temporal trends of publications, main authors, main countries of origin, main journals, main keywords, among others. For the development of the study, the technique of bibliometrics was used. Specifically, the R software was used to join the databases and the M.S. Excel and the Bibliometrix software. As a result, the field of studies developed slowly until the 2000s, with a growing trend from 2008 onwards, possibly associated with the subprime crisis. Subsequently, there were increases from 2016, with a peak of publications in 2022, possibly associated with the crisis caused by COVID-19. The most relevant study was that of Altman (1968). The main country for the field of studies in volume of citations was the United States, but China had the highest volume of publications. The most relevant journal was *Expert Systems With Application*, and when analyzing the keywords, the main ones were *classification, financial ratios, risk, neural-networks, models, discriminant-analysis, model, performance and prediction*. Trending topics were computational like *big data, machine learning, and deep learning*. Therefore, from the study, it was possible to carry out an important survey on the works involving the theme of risk and prediction of insolvency, using a broad set of data (*Scopus and Web of Science*).

Keywords: Bankruptcy prediction. Bankruptcy risk. Bibliometrics.

1 INTRODUÇÃO

A insolvência é um tema que tem atraído a atenção de pesquisadores ao longo do tempo (SHI; LI, 2019b). De acordo com Shi e Li (2019b) os estudos de previsão de insolvência cresceram nos últimos 50 anos e, diferentes métodos e variáveis têm sido usados para o desenvolvimento dos modelos preditivos, devido à aplicação em empresas de diferentes setores. Face à relevância da previsão de insolvência, os estudiosos se preocuparam principalmente com o desenvolvimento de modelos preditivos que fossem capazes de prever a situação de insolvência. De acordo com Prado *et al.* (2016), um dos trabalhos seminais e o mais influente sobre a temática foi o trabalho de Altman (1968). O estudo foi o primeiro a usar a análise multivariada como técnica de previsão.

Shi e Li (2019b) mencionaram a existência de importantes trabalhos de revisão sobre a insolvência. Os principais trabalhos de revisão tiveram como foco principal a revisão de estudos que utilizaram análises estatísticas para a construção dos modelos. Entre os trabalhos de revisão sobre a temática destacam-se trabalhos como os de Kumar e Ravi (2007), Balcaen e Ooghe (2006), Dimitras *et al.* (1996) e Bellovary, Giacomino e Akers (2007). Um ponto destacado por Shi e Li (2019b) foi que novas técnicas vêm sendo usadas, como as de aprendizado de máquina, com o intuito de que modelos cada vez mais robustos sejam desenvolvidos.

Face ao crescimento do número de trabalhos sobre a temática nos últimos anos (SHI; LI, 2019b), é relevante que novos trabalhos de revisão sejam desenvolvidos, para que se tenha conhecimento sobre como o campo de estudos vem se desenvolvendo. Apesar de tal relevância, de acordo com Prado *et al.* (2016), a análise de um grande número de trabalhos possui grande complexidade. Sendo assim, para otimizar tal análise, algumas técnicas sistemáticas de análise são utilizadas, como a bibliometria (PRADO *et al.*, 2016). A bibliometria é considerada como uma técnica quantitativa de análise que pode ser usada para medir a produção acadêmica, a propagação do conhecimento e verificação dos fluxos de informações sobre temas diversos do meio científico (VANTI, 2002).

Devido à importância da insolvência para os diversos agentes e devido ao desenvolvimento da temática nos últimos anos, este estudo analisará como o campo de previsão de insolvência se desenvolveu nos últimos anos. Sendo assim, o **objetivo** desse artigo foi de realizar uma revisão bibliométrica de literatura referente ao tema de risco e previsão de insolvência. Para alcançar o objetivo geral, foram desenvolvidos os seguintes passos: Realizar uma análise sistemática sobre o tema; Identificar tendências temporais de publicações,

principais autores, principais países de origem, principais periódicos, principais palavras-chave, entre outros.

A pesquisa foi relevante, pois a partir dela foi possível realizar um importante levantamento sobre os estudos envolvendo a temática do risco e previsão de insolvência, utilizando um amplo conjunto de dados, uma vez que foram usadas duas bases de dados (*Scopus* e *Web of Science*). Tal levantamento possibilita a identificação das principais tendências e também lacunas que poderão impulsionar estudos futuros sobre o tema em análise. Sendo assim, a partir desse levantamento foi possível aprofundar o conhecimento sobre o que já foi desenvolvido pela literatura e o mapeamento de autores, trabalhos, instituições, entre outros, que são relevantes para a temática.

2 O QUE SE TEM PESQUISADO SOBRE INSOLVÊNCIA E RISCO DE CRÉDITO

No campo das revisões sobre insolvência e risco de crédito, alguns trabalhos são mencionados como relevantes para a temática. Entre eles estão os trabalhos de Dimitras, Zanakis e Zopounidis (1996), Bellovary, Giacomino e Akers (2007), Kumar e Ravi (2007), Demyanyk e Hasan (2010) e Murcia, Murcia e Borba (2014). No que tange as revisões sistemáticas, cinco trabalhos principais são citados: Prado *et al.* (2016), Shi e Li (2019a), Shi e Li (2019b), Ubarhande e Chandani (2021) e Farias, Martínez e Martín-cervantes (2021).

No trabalho intitulado “*A survey of business failures with an emphasis on prediction methods and industrial*”, Dimitras, Zanakis e Zopounidis (1996) realizaram uma revisão de literatura sobre falência. Os autores identificaram que o principal método utilizado até então nos trabalhos foi a análise discriminante, seguida de trabalhos que utilizaram o modelo *logit*. Os autores identificaram também o surgimento de novos métodos de análise a partir de 1980, devidos às limitações da análise discriminante. Por fim, mencionou-se que as principais categorias de indicadores observadas nos trabalhos foram a de solvência (como ativo circulante sobre ativo e total de dívidas sobre o ativo total), e lucratividade.

Bellovary, Giacomino e Akers (2007) apresentam um resumo histórico acerca da temática de previsão de insolvência. Os autores ressaltaram inicialmente a existência de diferentes modelos de previsão de insolvência, bem como a utilização de novas técnicas de análise, principalmente devido aos avanços tecnológicos. Como resultado, os autores verificaram que o método que apresentou melhor faixa de precisão entre os trabalhos analisados foi o de redes neurais (71% a 100%). Adicionalmente, os autores mencionaram que o maior número de variáveis não garante maior capacidade preditiva do modelo.

O trabalho de Kumar e Ravi (2007) foi uma revisão de trabalhos publicados entre o período de 1968 a 2005. O trabalho enfatizou os tipos de técnicas adotados nos trabalhos analisados, mas outras dimensões também foram analisadas. Como resultado, os autores enfatizaram que na maioria dos trabalhos as empresas estiveram como objeto de estudo (não os bancos).

Ademais, os autores identificaram o uso de indicadores econômico-financeiros diversos, e que a maioria dos trabalhos analisou dados de organizações estadunidenses, seguidas por organizações de países europeus. Adicionalmente os autores observaram a tendência de substituição de técnicas estatísticas pela utilização da inteligência artificial, bem como a utilização de sistemas inteligentes híbridos, para a diminuição da limitação das técnicas individuais.

Demyanyk e Hasan (2010) realizaram uma revisão focando a previsão de crises financeiras e falências bancárias, enfatizando a crise de hipoteca do *subprime*, ocorrida nos Estados Unidos. A partir do trabalho, os autores verificaram que geralmente as técnicas estatísticas são utilizadas juntamente com técnicas de inteligência artificial, para a formulação de modelos com melhor desempenho. Os autores mencionaram que dentre as técnicas estatísticas para a previsão de falência bancária, a análise discriminante foi a mais usada por muito tempo, mas outras técnicas passaram a ser usadas, entre elas as de inteligência artificial, sendo que as redes neurais foram as mais utilizadas.

No trabalho de Murcia, Murcia e Borba (2014), os autores realizaram uma revisão de literatura focada na avaliação de crédito corporativo. A partir da análise, os autores verificaram que a maior parte dos trabalhos utilizou o modelo *logit* ordenado. Adicionalmente, as variáveis apontadas como mais relevantes do ponto de vista estatístico foram as de rentabilidade, tamanho, liquidez e endividamento. Os autores verificaram que apenas três dos 12 trabalhos analisados usaram dados de empresas brasileiras. Outro ponto de destaque foi a não concordância em relação as variáveis mais relevantes entre os trabalhos.

Dentre as revisões sistemáticas, a trabalho de Prado *et al.* (2016) introduziu a utilização de bibliometria nas revisões de literatura sobre o tema de insolvência e risco de crédito. A partir da utilização do método sistemático de bibliometria, os autores objetivaram realizar uma revisão para verificar e descrever as técnicas de análise multivariada utilizadas pela literatura.

De acordo com os autores, o trabalho mais relevante sobre a temática foi o de Altman (1968), seguido de Olshon (1980), que foram os mais citados e introduziram as técnicas de análise discriminante e regressão logística, respectivamente, como técnicas de previsão. Outro resultado destacado foi a identificação de diversas técnicas para os modelos preditivos, incluindo técnicas estatísticas e de aprendizado de máquina, sendo que as mais usadas foram redes neurais artificiais, regressão logística e análise discriminante.

Ademais, Prado *et al.* (2016) ressaltaram uma tendência de estudos de comparação de técnicas para a previsão. Houve também a identificação de um aumento no volume de citações e publicações após a crise de 2008, fatos que os autores consideram associados, uma vez que segundo eles essa crise mostrou que a análise de crédito possuía falhas.

Em relação aos principais periódicos, dois puderam ser destacados: *European Journal of Operational Research* e o *Expert Systems With Application*. Ao mencionar os países com os maiores volumes de citação, China, Estados Unidos e Taiwan foram os três principais. Os autores analisaram também as palavras que apareceram com maior frequência nos estudos, que foram, em ordem decrescente de volume de aparição: previsão de falência (109 vezes),

classificação (108 vezes) e análise discriminante (99 vezes), redes neurais (94 vezes), índices financeiros (76 vezes), pontuação de crédito (66 vezes), máquinas de vetores de suporte (60 vezes), falência (57 vezes), previsão (57 vezes e, por fim, modelos (56 vezes).

No trabalho de Shi e Li (2019a), os autores realizaram uma revisão sistemática de literatura sobre o tema de previsão de insolvência, considerando o período de 1968 a 2017. Corroborando os achados de Prado *et al.* (2016), os autores identificaram uma tendência de crescimento no volume de publicações após a crise de 2008, reforçando a importância do campo.

Ademais, os autores mencionaram que o estudo identificou a regressão logística e as redes neurais artificiais como as principais técnicas usadas no campo, não sendo, no entanto, as únicas, uma vez que diferentes técnicas também foram identificadas, incluindo de inteligência artificial e estatísticas. Por fim, os autores relataram a identificação de um baixo nível de co-autoria no campo de pesquisa, o que indica que há pouca colaboração entre os autores do campo, principalmente entre os mais influentes.

Já no trabalho de Shi e Li (2019b), os autores realizaram uma revisão bibliométrica enfatizando as pesquisas que utilizam técnicas de inteligência artificial para a elaboração de modelos preditivos. Segundo os autores, o primeiro estudo identificado sobre o tema que utilizou técnicas de inteligência artificial foi o de Tam (1991), intitulado “*Neural network models and the prediction of bank bankruptcy*”. Até o ano de 2008, os autores relataram a identificação de baixo nível de publicações. No entanto, corroborando os trabalhos de Prado *et al.* (2016) e Shi e Li (2019a), no período de 2008 a 2009 um aumento no volume de publicações foi observado, sendo que o interesse pelo tema apresentou crescimento desde então.

No estudo foram mencionados também os autores de maior peso de acordo com os volumes de publicação, sendo eles Li e Sun. Em relação à co-autoria por países, aqueles com maiores números de publicações foram China, Estados Unidos, Taiwan, Coréia do Sul e Espanha. A partir da análise de co-citação, verificou-se que os trabalhos mais citados foram Altman (1968) e Ohlson (1980), corroborando Prado *et al.* (2016), seguidos por Beaver (1966).

As palavras mais usadas foram previsão de falências, redes neurais, análise discriminante, previsão de dificuldades financeiras e classificação. Os autores verificaram também que as técnicas de redes neurais artificiais e análise discriminante foram as mais usadas por muito tempo. Por fim, ressalta-se que a métrica de árvore de decisão tem sido usada recentemente, mas alguns métodos como *fuzzy*, *data mining*, etc., têm sido pouco usados.

Outro trabalho de revisão sobre a temática foi o de Ubarhande e Chandani (2021). O estudo teve como objetivo revisar o conhecimento sobre classificação de crédito nas duas

últimas décadas, considerando o período de 2001 a 2020. Corroborando os resultados dos trabalhos de revisão anteriores, os autores identificaram um aumento no volume de estudos no período de 2008 a 2016.

Ubarhande e Chandani (2021) ressaltaram também que a preocupação principal dos estudos foi no desenvolvimento de métodos para classificação de crédito, sem uma análise mais aprofundada. Com os avanços tecnológicos das últimas décadas, os autores identificaram a popularidade e aceitação de técnicas como regressão, inteligência artificial e aprendizado de máquina. Outro resultado destacado foi o menor número de estudos de classificação de crédito relacionado a setores específicos, sendo que os principais foram setor financeiro/bancário e pequenas e médias empresas, sendo necessário a ampliação de estudos para a análise de setores específicos (UBARHANDE; CHANDANI, 2021)

Outro ponto relevante identificado por Ubarhande e Chandani (2021) foi que os estudos imediatamente posteriores à crise de 2008 tiveram como objetivo principal o desenvolvimento de novos métodos de classificação. Segundo os autores, uma maior atenção ao melhoramento de metodologias já existentes deve ser dada, uma vez que o melhoramento pode ser uma alternativa mais eficiente que a criação de novas metodologias. Por fim, os autores mencionam uma importante conclusão do estudo, que diz respeito à necessidade de desenvolvimento de uma metodologia de classificação híbrida, ou seja, que integre diferentes técnicas de previsão, como regressão, aprendizado de máquina, máquinas de vetor de suporte, entre outras.

No estudo de Farias, Martínez e Martín-cervantes (2021), os autores realizaram uma revisão bibliométrica sobre falhas de negócios. Os autores consideram que as falhas empresariais ocorrem quando uma empresa não consegue quitar suas obrigações, ou seja, quando se torna insolvente. Por meio do estudo, os autores identificaram que a maioria dos estudos foram publicados a partir do ano de 2006, sendo que o aumento do volume coincidiu com a crise financeira de 2008, podendo também estar ligado ao desenvolvimento de novas técnicas de previsão, segundo os autores.

Foi identificado também um aumento no volume de publicações a partir de 2016, fato que, segundo os autores pode estar associado a adoção dos Objetivos de Desenvolvimento Sustentável (ODS) implantados em 2015. Tal associação foi feita porque para a erradicação da pobreza e desenvolvimento sustentável é necessário o combate de falhas empresariais. Os autores identificaram também que as áreas principais que desenvolveram a temática foram negócios, gestão e contabilidade, responsáveis por 33,5% do total, em seguida estiveram economia, econometria e finanças que representaram 17,6% do total, e ciências sociais com 12,2%.

Ao analisar os principais periódicos, os que apresentaram maiores volumes de publicações sobre fracassos empresariais foram *Expert Systems with Applications*, *Journal of Business Research* e *Journal of Business Venturing*. Ao analisar os principais autores, os que apresentaram os maiores volumes de publicações foram Hui (21 artigos publicados) e Jie Sun (18 artigos publicados).

Ao considerar as instituições relevantes para a temática, os autores verificaram que a que apresentou o maior volume de trabalhos foi a “*Zhejiang Normal University*” (China) com 19 estudos publicados e 545 citações. A de maior impacto, no entanto, foi a “*Technical University of Crete*”, que apresentou 871 citações, com uma média de 124,43 citações por artigo. Por fim, os principais países foram os Estados Unidos, seguidos por Reino Unido China.

3 METODOLOGIA

O presente estudo caracteriza-se como uma análise bibliométrica. A bibliometria é entendida como uma técnica de análise quantitativa, que é utilizada para verificação da produtividade e da qualidade de pesquisas, através da utilização de medidas relacionadas ao número de publicações e citações de pesquisadores de determinada área (VANTI, 2002). É possível também fazer análises como de tendências dentro de uma área de pesquisa, analisar seu crescimento, verificar o surgimento de novos temas, identificar os principais autores, países e revistas, identificar processos de citação e co-citação, entre outros (VANTI, 2002).

De maneira complementar, Pinto *et al.* (2014) mencionam que o objetivo de estudos bibliométricos consiste no mapeamento intelectual sobre o conhecimento de determinado campo de estudos. Segundo os autores, esse tipo de análise tem sido utilizada no exame de diferentes temas dentro da área de negócios internacionais. Entre as análises proporcionadas pelo método estão: identificação de redes intelectuais, tendências de determinada área, principais autores e instituições, entre outras.

Uma vez que muitos estudos realizam análises bibliométricas considerando apenas uma base de dados (ECHCHAKOUI, 2020), neste estudo optou-se pelo uso de duas bases de dados. As bases de dados escolhidas foram Scopus e Web of Science, que são importantes bases de que reúnem diversos estudos de diferentes áreas de pesquisa. De acordo com Echchakoui (2020), o uso de ambas se faz mais adequado do que o uso de apenas uma delas, uma vez que o autor afirmou que elas são complementares.

Para a análise do conhecimento e de tendências sobre a temática de insolvência empresarial o *software* chamado *bibliometrix* foi usado, a partir do qual foi possível analisar a literatura e a evolução de publicações, co-autorias, co-citações, entre outros (SHI; LI, 2019a).

Prado *et al.* (2016) mencionaram que a análise bibliométrica possui limitações como erros de polissemia, visão geral da área e erros na etapa de coleta de dados. Para a correção de tais erros, Prado *et al.* (2016) utilizaram uma série de procedimentos realizados para a diminuição de erros e vieses. Desta maneira, o presente estudo apresenta também estágios para sua elaboração, adaptados a partir do trabalho de Prado *et al.* (2016), assim como de outros trabalhos que adotaram o método PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analysis*), introduzido por Moher *et al.* (2010), que de acordo com Echchakoui (2020) compreende as etapas de identificação de estudos, triagem, elegibilidade e inclusão. O Quadro 2 mostra os estágios usados para a construção do estudo.

Quadro 2 — Estágios de elaboração do estudo.

Estágio	Procedimento	Descrição
1: operação de pesquisa	1.1	Escolha da base científica de dados.
	1.2	Delimitação de termos que representam o campo.
2: Procedimentos de pesquisa (filtros)	2.1	Título (termos do campo).
	2.2	Uso de <i>underline</i> (expressão exata) e asterisco.
	2.3	Buscas individuais de cada termo.
	2.4	Retirada de termos não relacionados ao campo e adaptações.
	2.5	Nova busca a partir da adaptação da <i>string</i> .
	2.6	Buscas individuais de cada termo após adaptação.
	2.7	Filtro 1: Delimitações de artigos e revisões.
	2.8	Filtro 2: Todos os anos.
	2.9	Filtro 3: Todas as áreas.
	2.10	Filtro 4: idioma = inglês.
3: Procedimentos de seleção: Banco de dados	3.1	<i>Download</i> de referências - <i>software Mendeley</i> .
	3.2	<i>Download</i> de referências em formato de planilha.
	3.3	<i>Download</i> de referências para uso nos softwares.
	3.4	Organização de referências no <i>Mendeley</i> .
	3.5	Análise de matriz na organização de planilhas.
	3.6	Importação de dados para o <i>software</i> de análise.
4: Adequação e organização dos dados	4.1	Eliminação de artigos duplicados.
	4.2	Junção das bases de dados <i>Scopus</i> e <i>Web of Science</i> por meio do R.
5: análise da produção científica	5.1	Análise de tendências temporais e de volumes de publicação.
	5.2	Análise de referências e artigos mais citados.
	5.3	Análise do país de origem.
	5.4	Análise de periódicos mais citados.
	5.5	Análise de autores citados e co-citados.
	5.6	Análise de palavras-chave.
	5.7	Estudo de tendências.

Fonte: Adaptado de Prado *et al.* (2016).

Desta maneira, quatro estágios principais serão seguidos para a elaboração do estudo. Cada estágio compreenderá um conjunto de procedimentos para a realização das análises, que serão detalhados a seguir. O intuito do uso desses procedimentos foi evitar a incidência de erros na pesquisa.

3.1 Estágio 1: operacionalização da pesquisa

Para a elaboração do presente estudo, optou-se pela utilização de duas bases de dados para a realização da consulta sobre o tema. As bases de dados escolhidas foram a *Scopus* e a *Web of Science*. Optou-se pelo uso de duas bases de dados, uma vez que, de acordo com Echchakoui (2020), o uso das duas bases de dados fornece uma visão ampla do conhecimento desenvolvido no campo de estudos, o que se torna limitado caso apenas uma das bases de dados seja usada. Ademais, menciona-se que essas bases de dados são consideradas como as mais

relevantes e confiáveis para a realização de análises bibliométricas (DE LA HOZ-CORREA; MUÑOZ-LEIVA; BAKUCZ, 2018; KNANI; ECHCHAKOUI; LADHARI, 2022; MAIA et al., 2019), tornando-as apropriadas para as análises propostas no estudo .

A delimitação dos termos utilizados na pesquisa, se baseou em artigos seminais e relevantes da temática, bem como em artigos de revisão sistemática sobre a temática de insolvência realizados anteriormente. Nos títulos dos trabalhos de Altman (1968) e de Ohlson (1980) foram identificados os termos “*Bankruptcy*” e “*Prediction*”. Com base inicialmente nesses dois estudos, nas buscas foram utilizados os termos “*Bankruptcy*” e “*Bankruptcy prediction*”. Outro termo usado na busca foi o de “*Business failure*”, identificado nos títulos dos trabalhos de Dimitras *et al.*, (1999) e Dimitras, Zanakis e Zopounidis,(1996). O termo “*Credit scoring*”, por sua vez, esteve presente no título do trabalho de West (2000).

Outros termos usados em revisões anteriores também foram considerados, como “*Credit risk*”, usado no trabalho de Prado *et al.* (2016); e “*Default firm*”, “*early-warning*”, “*failure prediction*”, “*financial distress*” e “*insolvency*”, do trabalho de Shi e Li (2019a). Estarão na busca os termos “*default prediction*”, “*financial failure*”, “*financial distress*”, “*insolvency*” e “*business failure*”, do trabalho de Shi e Li (2019b). Por fim, foram consideradas também algumas palavras adicionais ou variações de algumas sugeridas pelos estudos, que estarão detalhadas na *string* de buscas.

3.2 Estágio 2: procedimentos de pesquisa

Conforme mencionado no estágio 1, a primeira *string* de busca foi definida a partir dos trabalhos de Prado *et al.* (2016), Shi e Li (2019a) e Shi e Li (2019b), considerando, conforme já mencionado, palavras presentes nas buscas de revisões anteriores e algumas variações e também palavras adicionais relacionadas ao campo pesquisado. Após a filtragem de termos que não retornaram estudos do campo de pesquisa e da realização de algumas adaptações a alguns termos definiu-se uma a *string* de buscas, que apresentou inicialmente 3.015 trabalhos na *Scopus* e 2.342 trabalhos na *Web of Science*. A data da pesquisa foi 06 de Janeiro de 2023, logo, os artigos publicados até essa data foram considerados na análise.

A partir dos resultados dessa busca novas filtrações foram realizadas. Inicialmente, filtrou-se pelo tipo de documento, uma vez que essa filtragem de acordo com Prado et al. (2016) se faz necessária para a realização das análises nos *softwares* bibliométricos. Sendo assim, foram mantidos apenas os documentos classificados como artigo ou revisão, todos os outros

tipos foram selecionados e excluídos. Após essa exclusão, chegou-se a um total de 1.391 trabalhos na *Web of Science* e 1.995 na *Scopus*.

Após isso, foram excluídos os trabalhos que não estavam em inglês, para uma maior padronização da amostra (BENOMAR et al., 2022; ECHCHAKOUI, 2020; ECHCHAKOUI; BARKA, 2020; KNANI; ECHCHAKOUI; LADHARI, 2022). A partir dessa filtragem, o total final de trabalhos na *Web of Science* foi de 1.362 e na *Scopus* de 1.905. A Figura 5 mostra a *string* final de busca que foi utilizada na *Web of Science*.

Figura 5 — *String* final *Web of science*.

Título: (bankruptc* NEAR/2 predict* OR
bankruptc* NEAR/2 forecast* OR
insolvenc* NEAR/2 predict* OR
default_predict* OR
business_failure* OR
credit_scor* OR credit_risk_model* OR
credit_risk_assessment OR
credit_risk_classificat* OR
credit_risk_predict* OR
credit_risk_forecast*).

Fonte: Da autora (2023).

A Figura 6 mostra a *string* de buscas usada na base de dados *Scopus*.

Figura 6 — *String* final *Scopus*.

Título: ((bankruptc* W/2 predict*) OR
(bankruptc* W/2 forecast*) OR
(insolvenc* W/2 predict*) OR
(default_predict*) OR
(business_failure*) OR
(credit_scor*) OR (credit_risk_model*) OR
(credit_risk_assessment) OR
(credit_risk_classificat*) OR
(credit_risk_predict*) OR
(credit_risk_forecast*)).

Fonte: Da autora (2023).

Uma vez que as bases de dados apresentam trabalhos duplicados, realizou-se por meio do R a mesclagem das bases de dados *Scopus* e *Web of Science*. Após as filtrações mencionadas, os arquivos com as informações necessárias para a análise bibliométrica e que formaram a base de dados final foram baixados no formato BibTex (*Scopus*) e Txt (*Web of science*). Após isso, as bases de dados foram combinadas por meio do Rstudio, onde os trabalhos duplicados também foram excluídos das amostras. Após a combinação das bases, o total de trabalhos foi de 2.066. Para considerar apenas estudos de fontes reconhecidas, os trabalhos que não

apresentaram o *International Standard Serial Numbers* (ISSN) foram excluídos da amostra (BENOMAR et al., 2022; ECHCHAKOUI, 2020; ECHCHAKOUI; BARKA, 2020; KNANI; ECHCHAKOUI; LADHARI, 2022). Dessa maneira, a amostra final do estudo foi composta por 1.995 trabalhos.

3.3 Estágio 3: Procedimentos de seleção

As referências e referências citadas pelos trabalhos da amostra foram exportadas para o gerenciador de referências *Mendeley* e foram baixados arquivos em formato de planilha e de texto, para as realizações das análises através do auxílio do M.S. *Excel* e do *software VOSviewer* (VAN ECK; WALTMAN, 2010).

3.4 Estágio 4: análise da produção científica:

Para a realização das análises, foram utilizados o M.S. *Excel 2007* e o *software Bibliometrix*. A partir dessas ferramentas foi possível a realização do mapeamento do campo de estudo, apresentando os trabalhos e periódicos mais citados, países com maior número de publicações, redes de co-autoria e cocitação, a evolução de publicações sobre o tema ao longo dos anos, principais palavras-chave utilizadas, entre outros.

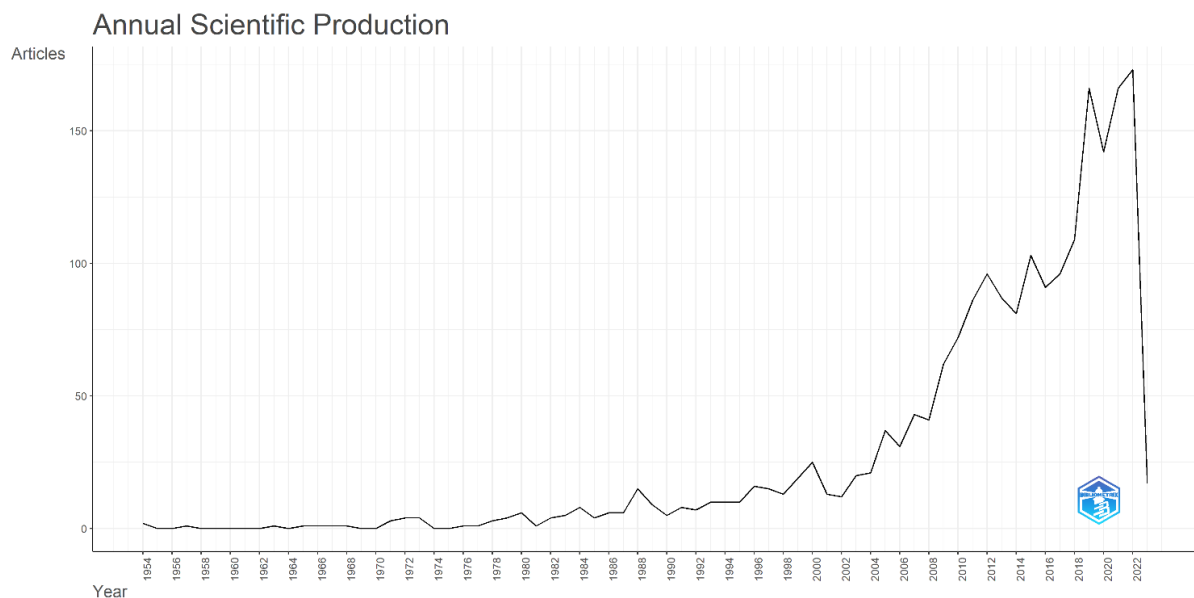
4 RESULTADOS

Nesta Seção os resultados da análise bibliométrica foram detalhados.

4.1 Tendências temporais e análises de volume de publicação

O ano inicial com estudos publicados sobre a temática foi o ano de 1954. Nesse ano, dois trabalhos foram publicados, sendo eles o de Marburg (1954) intitulado “*A Study of Small Business Failure Smith Griggs of Waterbury*” e o trabalho de Lomax (1954), intitulado *Business Failures - Another Example Of The Analysis Of Failure Data*. A partir de então, o campo de pesquisas se desenvolveu lentamente, com baixos volumes de publicações nas décadas seguintes, apresentando aumento somente após o ano 2000, conforme ilustrado pela Figura 7.

Figura 7 — Tendências temporais.



Fonte: Da autora (2023).

Conforme observado na Figura 7, é possível identificar um aumento de publicações a partir do ano de 2008, resultado também identificado por Prado *et al.* (2016) e Shi e Li (2019a), provavelmente associado à crise econômica do *subprime*. Percebeu-se também outro aumento das publicações a partir de 2016, resultado que corrobora o de Farias, Martínez e Martín-cervantes (2021), que mencionaram que este pode ser um reflexo da adoção dos Objetivos de Desenvolvimento Sustentável (ODS), que foram implantados em 2015. Essa associação, se deve ao fato de que para atingir esses objetivos, as falhas empresariais devem ser minimizadas.

Essa tendência de crescimento continuou, sendo que o auge das publicações foi em 2022, com um total de 172 trabalhos publicados, podendo ser um possível reflexo da crise causada pela pandemia de COVID-19. Como a pesquisa foi realizada no ano de 2023, artigos publicados nesse ano já compuseram a amostra, até o dia da coleta de dados (06 de Janeiro de 2023) haviam 17 artigos publicados sobre a temática.

4.2 Análise de artigos mais citados

A elaboração do *ranking* de trabalhos mais relevantes sobre o campo de pesquisa se deu a partir dos trabalhos mais citados das bases de dados, conforme apresentado a Tabela 1.

Tabela 1 — Artigos com maiores volumes de citações.

Título	Autores	Título Da Fonte	Ano	Citações	Média ano
<i>Financial ratios, discriminant analysis and prediction of corporate bankruptcy</i>	Altman, E.I.	<i>Journal of Finance</i>	1968	5.095	93
<i>Financial ratios and the probabilistic prediction of bankruptcy</i>	Ohlson, J.A	<i>Journal of Accounting Research</i>	1980	2.106	49
<i>A rasch model for partial credit scoring</i>	Masters, G.N.	<i>Psychometrika</i>	1982	2.090	51
<i>Forecasting bankruptcy more accurately: A simple hazard model</i>	Shumway, T	<i>Journal of Business</i>	2001	1.008	46
<i>Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review</i>	Kumar, P. R.; Ravi, V.	<i>European Journal of Operational Research</i>	2007	596	37
<i>Neural network credit scoring models</i>	West, D.	<i>Computers & Operations Research</i>	2000	569	25
<i>Learning from business failure: Propositions of grief recovery for the self-employed</i>	Shepherd, D.A	<i>Academy of Management Review</i>	2003	516	26
<i>Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters</i>	Min, J.H.; Lee, V.C	<i>Expert Systems with Applications</i>	2005	507	28
<i>Credit Scoring With a Data Mining Approach Based on Support Vector Machines</i>	Huang, Z; Chen, H.C.; Hsu,C.J.; Chen, W.H.; Wu, S.S.	<i>Expert Systems with Applications</i>	2007	467	29
<i>Benchmarking state-of-the-art classification algorithms for credit scoring</i>	Baesens, B; Van Gestel, T; Viaene, Ss; Stepanova, M; Suykens, J; Vanthienen, J.	<i>Journal of the Operational Research Society</i>	2003	467	23

Fonte: Da autora (2023).

Por meio da Tabela 1 foi possível identificar os trabalhos mais relevantes do campo de estudos. Consistente com os resultados de Prado *et al.* (2016) e Shi e Li (2019b) o trabalho mais relevante, de acordo com o volume de citações foi o de Altman (1968), com 5.095 citações no período, com uma média aproximada de 93 citações por ano. O trabalho de Altman (1968), “*Financial ratios, discriminant analysis and prediction of corporate bankruptcy*”, foi o primeiro a utilizar análise multivariada para a previsão de insolvência (análise discriminante), sendo considerado um dos estudos seminais do campo de estudos.

Em seguida esteve o trabalho de Ohlson (1980), nomeado “*Financial ratios and the probabilistic prediction of bankruptcy*”, resultado que também corrobora o trabalho de Prado *et al.* (2016). Com um total de 2.106 citações e média anual de aproximadamente 49, este estudo foi o primeiro a utilizar a regressão logística para a previsão de insolvência, sendo também importante para o desenvolvimento do campo de estudos.

Em terceiro esteve o trabalho “*A rasch model for partial credit scoring*”, de Masters (1982). Com 2.090 citações e média anual de 51, o estudo desenvolveu um modelo de “crédito parcial”, que possui como característica de separabilidade de parâmetros, que de acordo com o autor possibilita comparações objetivas de diferentes indivíduos. Assim, os parâmetros de cada indivíduo são estimados e particularidades de cada indivíduo são obtidas, independentemente das características da amostra.

Como quarto mais citado esteve o trabalho de Shumway (2001), intitulado “*Forecasting bankruptcy more accurately: A simple hazard model*”, com 1.008 citações e média anual aproximada de 46 citações.

Em quinto lugar esteve o trabalho de Kumar e Ravi (2007) “*Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review*”, com 596 citações, com média anual de 37 aproximadamente.

Em sexto lugar, com 569 citações esteve o trabalho “*Neural network credit scoring models*”, de West (2000), apresentando em média 25 citações anuais.

O sétimo trabalho mais citado foi o intitulado “*Learning from business failure: Propositions of grief recovery for the self-employed*”, de Shepherd (2003), com 516 citações e média anual de aproximadamente 26 citações.

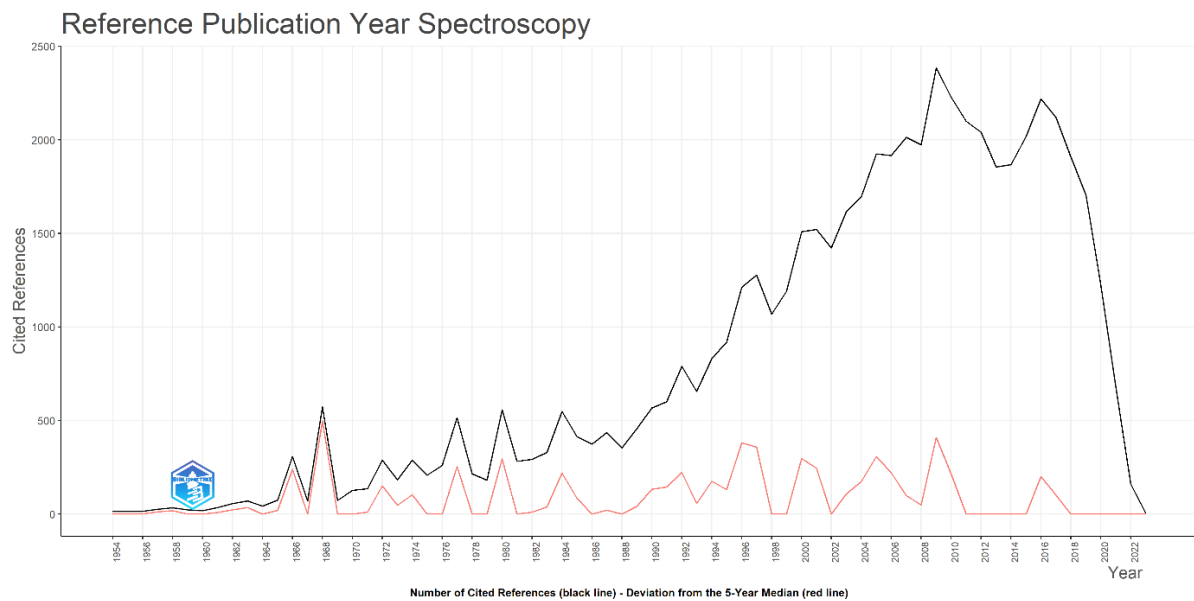
Em oitavo lugar esteve o trabalho de Min e Lee (2005), “*Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters*”, com 507 citações, e média anual de aproximadamente 28.

Em nono, esteve o trabalho de Huang, Chen e Wang (2007) intitulado “*Credit Scoring With a Data Mining Approach Based on Support Vector Machines*”, com 467 citações e média anual aproximada de 29.

Por fim, em décimo lugar esteve o trabalho “*Benchmarking state-of-the-art classification algorithms for credit scoring*”, de Baesens *et al.* (2007), com 467 citações e média anual de aproximadamente 23.

Por meio da Figura 8, é possível verificar os trabalhos que podem ser considerados como as origens históricas do campo de estudos, revelando os anos em que as descobertas relevantes para o campo foram realizadas (MARX *et al.*, 2014). Assim, por meio dessa análise, é possível identificar estudos relevantes para o campo. As descobertas ou estudos relevantes para o campo são representados pelos picos da linha vermelha da Figura 8.

Figura 8 — Referência publicação ano espectroscopia.



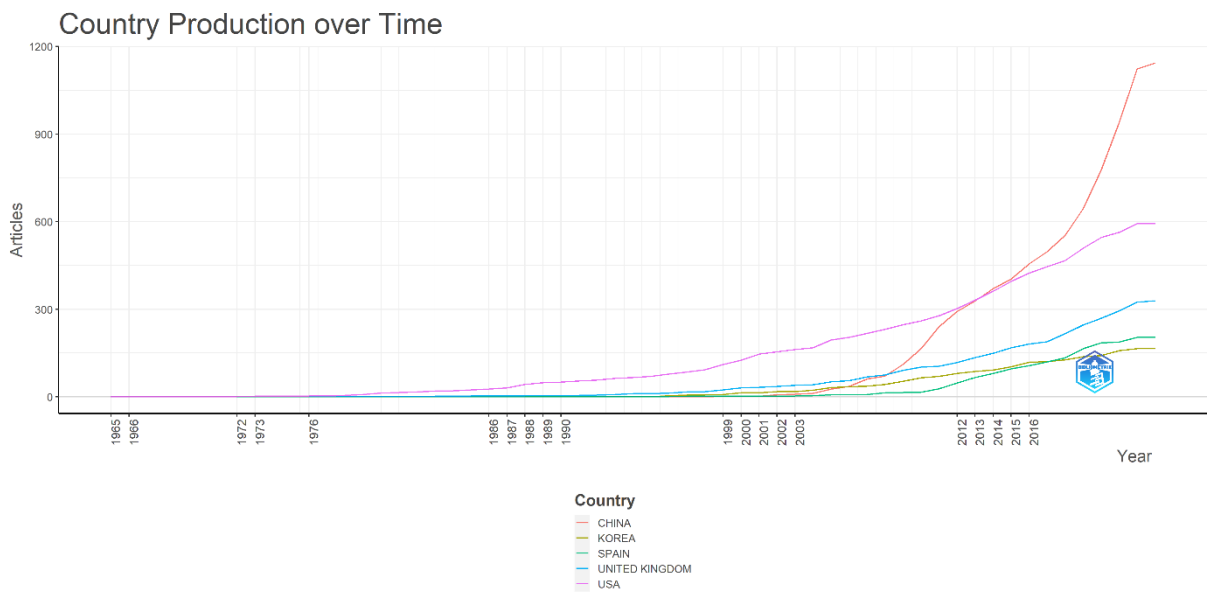
Fonte: Da autora (2023).

Por meio da Figura 8 percebe-se que os trabalhos que representam esses picos foram os de Altman (1968), Olshon (1980) e Beaver (1966). Esses resultados confirmam a influência dos estudos de Altman e Olshon para o campo de estudos, que introduziram a análise multivariada (análise discriminante) e a regressão logística como técnicas de previsão, respectivamente. O estudo de Beaver também pode ser considerado como relevante para o campo, uma vez que foi um dos estudos seminais, que utilizou a análise univariada para a previsão de insolvência. Ademais, mencionam-se também os estudos de West (2000) e Zmijewski (1984), que também apresentaram picos, sendo também estudos relevantes para o campo.

4.3 Análise do país de origem

A partir de análises do país de origem, foi possível identificar os países que mais contribuíram para o desenvolvimento do campo de estudos. Os países com maiores frequências de publicação foram China (1.142), Estados Unidos (589) e Reino Unido (327). A Figura 9 mostra a evolução dos volumes de citações por países ao longo dos anos:

Figura 9 — Evolução do volume de publicações por país.



Fonte: Da autora (2023).

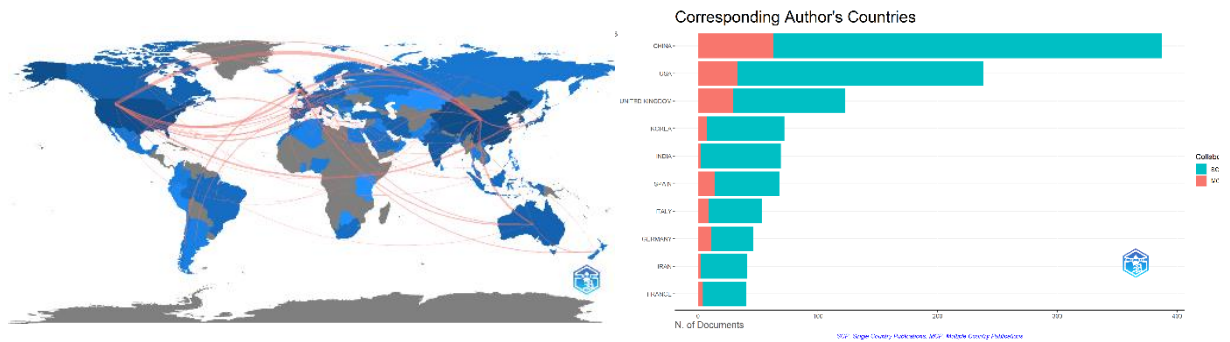
Por meio da Figura 9, percebe-se que a China ultrapassou os Estados Unidos a partir do ano de 2012, apresentando crescimento acentuado das publicações nos últimos 10 anos. Percebe-se também que o aumento de publicações dos Estados Unidos, assim como do Reino Unido se deram principalmente após os anos 2000, o que mostra que o campo de estudos se desenvolveu principalmente nos últimos 20 anos.

A Figura 10 mostra o mapa de colaboração entre os países, que têm como países principais a China, Estados Unidos e o Reino Unido. Ademais, mostra também os países dos autores correspondentes, que indicam a quantidade de trabalhos que foram desenvolvidos por autores do mesmo país (SCP) e a quantidade de trabalhos em que houve a colaboração de autores de países diferentes (MCP).

A China apresentou um total de 387 trabalhos, sendo 327 com autores apenas do país e 63 trabalhos com colaboração. Em seguida esteve os Estados Unidos, com 238 trabalhos, com 205 com autores do país e 33 trabalhos com colaboração de autores de outros países. Em terceiro

esteve o Reino Unido, com 123 estudos, sendo 94 com autores do país e 29 trabalhos com colaboração.

Figura 10 — Mapa de colaboração entre os países e países do autor correspondente.



Nota: SCP = Single Country publication (trabalhos apenas com autores do país).
MCP = Multiple Countries publication (trabalhos com multi-coautoria entre países).
Fonte: Da autora (2023).

Por meio da Figura 10, percebe-se que os países com maiores volumes de colaboração foram de China para Estados Unidos (31), de Estados Unidos para Reino Unido (15) e de China para Reino Unido (13). Isso mostra que os países mais relevantes da temática também apresentaram colaborações entre si, fato que pode contribuir para a relevância dos estudos, bem como para a produção e disseminação de conhecimento sobre o campo de estudos.

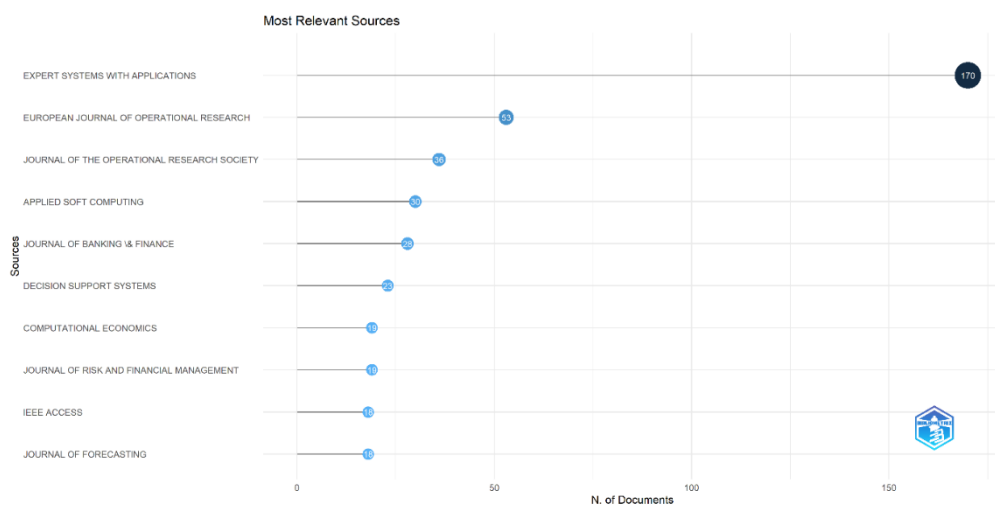
Ao se tratar do volume de citações, no entanto, os Estados Unidos ainda lideram com volume de citações de 16.576, seguido de China com 10.556 e Reino Unido, com 5.566. Tal fato confirma a relevância dos três países para o desenvolvimento do campo de estudos. Ademais, esses resultados corroboram parcialmente os achados de Prado *et al.* (2016), uma vez que os três países com maiores volumes de citações foram China, Estados Unidos e Taiwan. Esses resultados também corroboram parcialmente os achados de Farias, Martínez e Martín-cervantes (2021), que apresentaram Estados Unidos, Reino Unido e China como os principais países.

Além disso, ressalta-se que os Estados Unidos iniciaram suas produções no campo de estudos antes da China, e teve um crescimento até os anos 2000, mas se manteve como país com maiores volumes de publicação até 2012, conforme mostrado na Figura 9. Tal fato pode ser o responsável por manter o país com líder entre os volumes de citações, mesmo a China tendo-o superado em volume de trabalhos na última década.

4.4 Análise dos periódicos mais citados

Por meio da análise de periódicos mais citados foi possível identificar os que mais contribuíram para o desenvolvimento do campo de estudos. No que se refere ao número de documentos publicados, o principal periódico foi o “*Expert Systems With Applications*”, resultado corrobora o trabalho de Farias, Martínez e Martín-cervantes (2021). O número de trabalhos publicados foi de 170, conforme a Figura 11 ilustra.

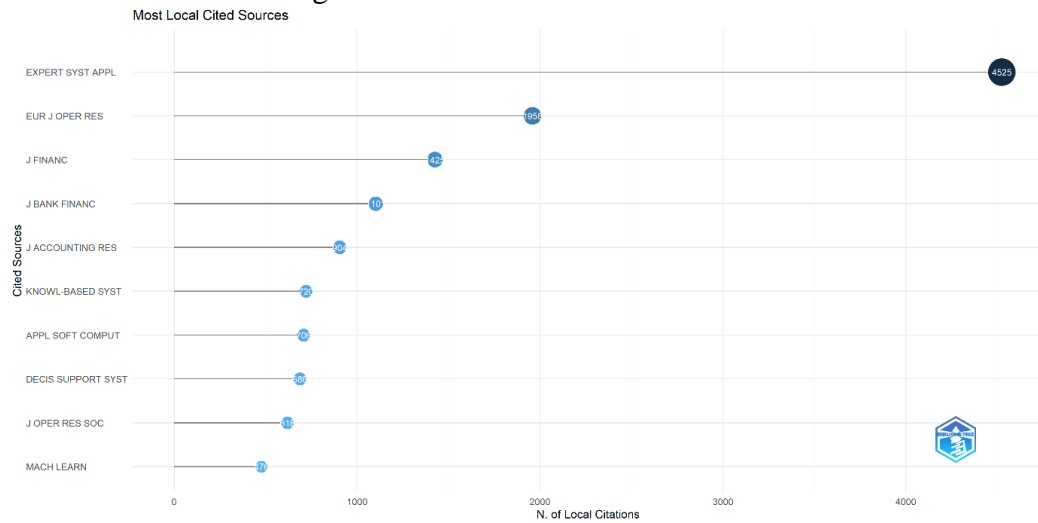
Figura 11 — Periódicos com maiores volumes de publicações.



Fonte: Da autora (2023).

Em seguida estiveram o “*European Journal of Operational Research*”, com 53 publicações e o “*Journal of the Operational Research Society*”, com 26 publicações. Ao se tratar do volume de citações, o “*Expert Systems With Applications*” e o “*European Journal of Operational Research*” também aparecem em primeiro e segundo lugar, com 4.525 e 1.958 citações, respectivamente, conforme mostrado na Figura 12.

Figura 12 — Periódicos mais citados.



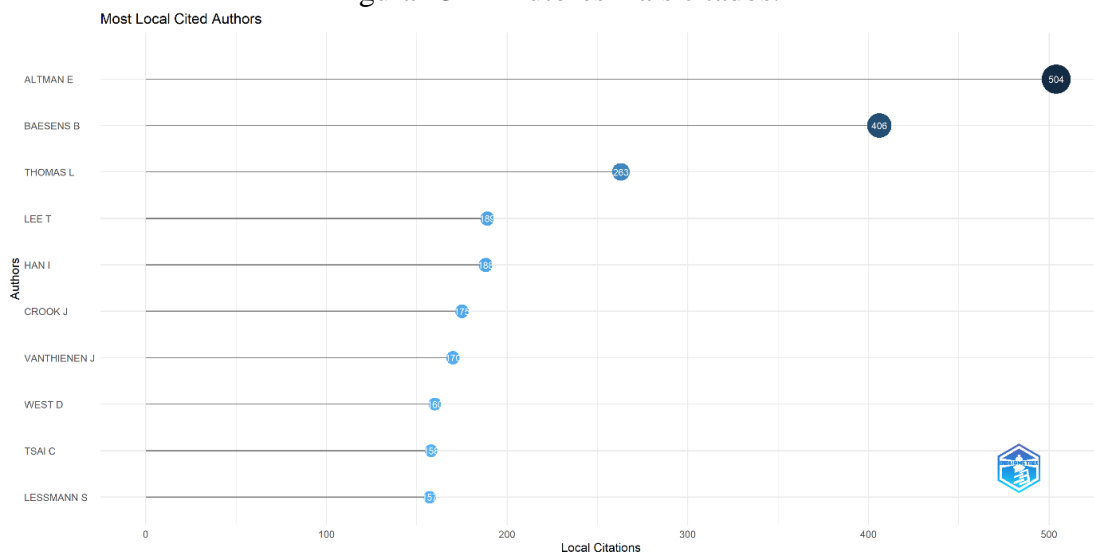
Fonte: Da autora (2023).

Em terceiro lugar, no entanto, apareceu o “*Journal of Finance*” (1.425 citações), conceituado periódico da área de finanças.

4.5 Análises de autores Citados

A partir da análise de autores mais citados é possível identificar os principais nomes do campo de estudos analisado. Por meio da Figura 13, percebe-se que o autor mais citado foi Altman, com 504 citações, resultado que corrobora os estudos de Prado *et al.* (2016) e Shi e Li (2019b). O trabalho de Altman, como já mencionado, foi um dos seminais do campo de estudos, o qual introduziu a análise multivariada como um método de previsão.

Figura 13 — Autores mais citados.

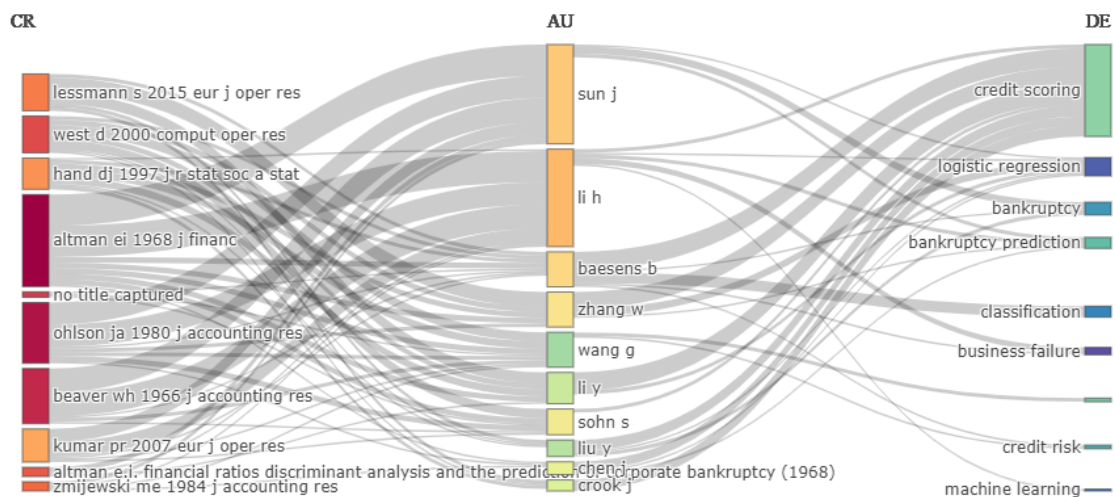


Fonte: Da autora (2023).

Em segundo lugar esteve Baesens, com 406 citações e em quarto esteve Thomas, com 263 citações.

Apesar do volume de citações, ao analisar o volume de publicações, percebe-se que os autores com maiores volumes de trabalhos publicados no período foram Li e Sun, com 25 trabalhos cada. Esse resultado corrobora o estudo de Shi e Li (2019b). Na Figura 14 é possível verificar os autores com maiores volumes de trabalhos publicados sobre o campo de estudos (AU), bem como as raízes intelectuais (CR) e os conteúdos pesquisados (DE).

Figura 14 — Gráfico de três campos: volume de publicações por autor.



Fonte: Da autora (2023).

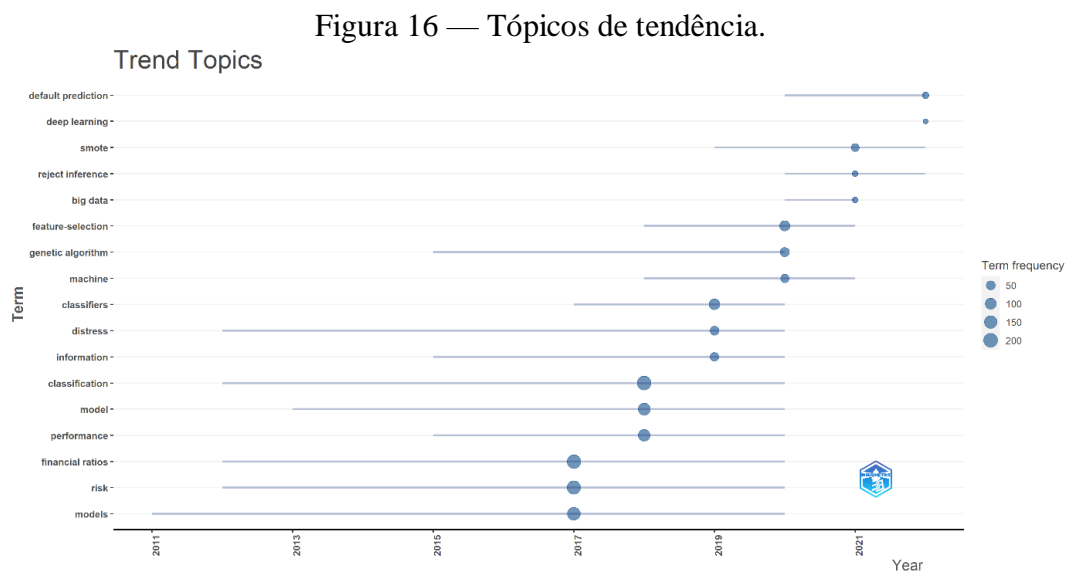
Conforme observado na Figura 14, o autor Sun desenvolveu estudos que tiveram como base o conhecimento gerado nos estudos de Altman (1968), Olshon (1980), Beaver (1966), Kumar (2007) e Zmijewski (1984), tendo como conteúdos pesquisados a falência (*bankruptcy*), previsão de falência (*bankruptcy prediction*) e a técnica de regressão logística (*logistic regression*).

Li também teve como base os estudos de Altman (1968), Olshon (1980), Beaver (1966), Kumar (2007) e Zmijewski (1984), mas os principais conteúdos pesquisados foram pontuação de crédito (*credit scoring*), regressão logística (*logistic regression*), previsão de falência (*bankruptcy prediction*), falência de negócios (*business failure*) e aprendizado de máquinas (*machine learning*).

Baesens apareceu como o segundo maior volume de publicações, com 15 trabalhos no total. No que se refere as raízes usadas pelo autor, estiveram os trabalhos de Lessman (2015), West (2000), Hand (1997), Altman (1968), Olshon (1980), Beaver (1966), Kumar (2007) e

models (163), *discriminant-analysis* (149), *model* (129), *performance* (119) e *prediction* (113). Tais resultados mostram que as principais técnicas de previsão tratadas foram a de redes neurais artificiais e análise discriminante. A partir da imagem nota-se também que outras técnicas também têm sido tratadas no campo de estudos, como a técnica de SVM (*support vector machine*).

A Figura 16 mostra os tópicos de tendência dos últimos anos, baseado nas palavras-chave dos estudos.

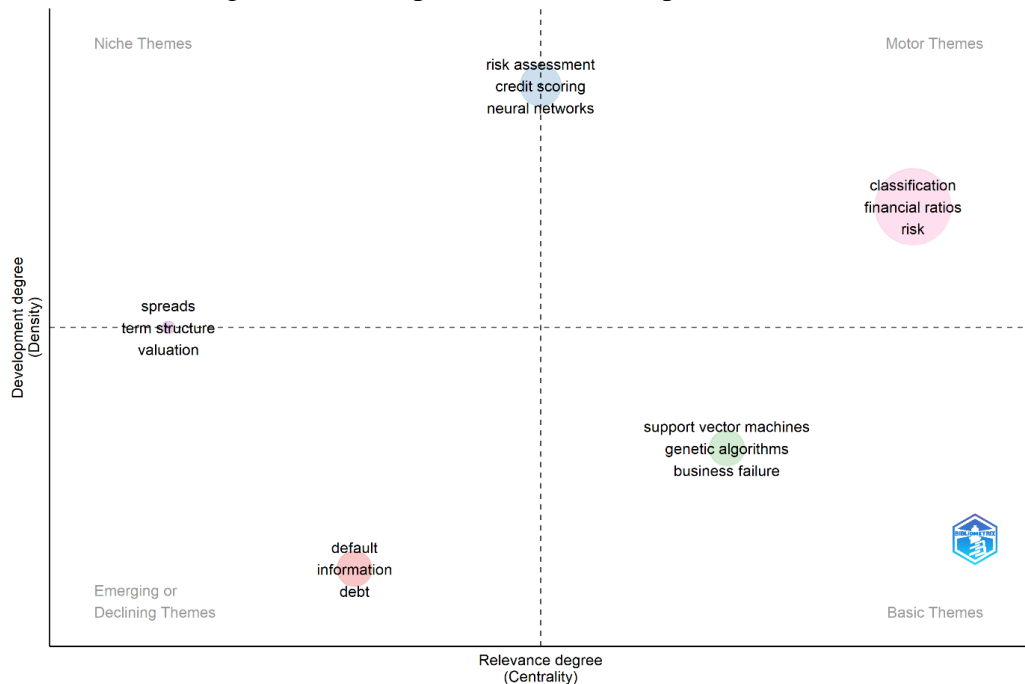


Fonte: Da autora (2023).

A partir da Figura 16, percebeu-se que os tópicos que têm sido retratados dentro do campo de estudos foram tópicos computacionais, como *big data*, *machine learning*, e *deep learning*, o que mostra que a introdução de algoritmos cada vez mais robustos têm sido uma tendência no campo de previsão e risco de insolvência.

Na Figura 17 é possível ver o mapa temático do campo, que mostra os principais temas do campo de estudos e a fase em que se encontram, baseado nos conceitos de centralidade (importância do tema) e densidade (desenvolvimento do tema) (COBO *et al.*, 2011).

Figura 17 — Mapa temático do campo de estudos.



Fonte: Da autora (2023).

Por meio de Figura 17, percebe-se no quadrante de temas motores (com alta centralidade e alta densidade) estão os temas classificação (*classification*), indicadores financeiros (*financial rates*) e risco (*risk*). Esses são considerados como temas estabelecidos e relevantes para o campo de estudos.

No quadrante de temas básicos (alta centralidade e baixa densidade) estão máquinas de vetores de suporte (*support vector machines*), algoritmos genéticos (*genetic algorithms*) e falência de negócios (*business failure*). Esses são temas relevantes para o campo, mas que ainda não foram tão desenvolvidos.

No quadrante de temas emergentes ou em declínio apareceram os termos *default*, informação e dívida. Esses são temas com baixa centralidade e baixa densidade, sendo eles emergentes ou em declínio dentro do campo de estudos.

Por fim, o quadrante de temas muito especializados (baixa centralidade e alta densidade) não apresentou-se bem definido. No entanto, os temas avaliação de risco (*risk assessment*), pontuação de crédito (*credit scoring*) e redes neurais artificiais (*neural networks*) estiveram na fronteira entre o quadrante de temas muito especializados e temas motores, indicando que esses temas estão caminhando para serem relevantes e bem desenvolvidos dentro do campo de estudos. Já os temas *spread*, estrutura a termo (*term structure*) e *valuation* estiveram na fronteira entre o quadrante de emergentes ou em declínio e temas muito especializados. Tal resultado

indica que esses são temas emergentes, que estão caminhando para um alto desenvolvimento, apesar da baixa centralidade até então.

5 CONSIDERAÇÕES FINAIS

O risco de crédito e a insolvência empresarial são fatores relevantes para diferentes agentes da sociedade. A antecipação do evento de insolvência beneficia empresas, fornecedores, investidores, entre outros, uma vez que tal previsão permite a realização de ações corretivas, evitando um possível risco sistêmico, decorrente de uma falência. Face à relevância do campo de estudos, este trabalho objetivou realizar uma revisão bibliométrica sobre previsão de insolvência e risco de crédito. Por meio do estudo foi possível realizar o mapeamento do campo.

Para isso, utilizou-se uma amostra composta por artigos presentes em duas bases de dados consideradas como relevantes e complementares, sendo elas a Scopus e Web Of Science. Por meio do Bibliometrix e do M.S. Excel (2010), foi possível a elaboração de análises da produção científica da área temática.

Por meio da análise de tendências de publicações, verificou-se que o ano inicial de estudos sobre o campo de pesquisas foi o ano de 1954, com dois trabalhos publicados, que foram o de Marburg (1954) intitulado “*A Study of Small Business Failure Smith Griggs of Waterbury*” e o de Lomax (1954), nomeado “*Business Failures - Another Example Of The Analysis Of Failure*”.

O campo de estudos se desenvolveu lentamente principalmente até os anos 2000, com uma tendência de crescimento principalmente a partir de 2008 para 2009, fato que deve ser um reflexo da crise do *subprime*, que também foi identificado por Prado *et al.* (2016). Posteriormente, houve aumentos consideráveis a partir do ano de 2016, possivelmente associados aos ODS, conforme Farias, Martínez e Martín-cervantes (2021), com auge de publicações no ano de 2022, com um total de 172 trabalhos, fato que pode ser um reflexo da crise decorrente da pandemia de COVID-19.

No que se refere ao volume de citações, o estudo de Altman (1968) foi considerado como mais influente do campo de estudos, sendo que a relevância do autor se confirmou por meio da análise de autores mais citados. No entanto, ao se tratar do volume de publicações, os autores Sun e Li apresentaram os maiores volumes, ambos com 25 estudos publicados.

A análise dos países de origem mostrou que os principais para o campo de estudos foram Estados Unidos, China e Reino Unido. Apesar do maior volume de publicações pertencer à China, o país com maior volume de citações foi os Estados Unidos, que iniciou o desenvolvimento de estudos antes da China.

Ao analisar os periódicos do campo de estudos, verificou-se que os mais relevantes foram o *Expert Systems With Application*, o *European Journal of operational Research*, *Journal of the Operational Research Society* e o *Journal of Finance*.

Ao analisar as palavras-chave do campo de estudo, foi possível identificar que as principais foram: *classification, financial ratios, risk, neural-networks, models, discriminant-analysis, model, performance* e *prediction*. Resultados que mostram que as principais técnicas de previsão tratadas até então foram as redes neurais artificiais e a análise discriminante. Ao analisar os tópicos de tendência, percebeu-se que principalmente nos últimos anos, tópicos computacionais como *big data, machine learning, e deep learning* estiveram presentes no campo de estudos, o que sugere que o uso de algoritmos mais robustos têm sido uma tendência no que se refere à previsão e ao risco de insolvência.

Por tudo isso, foi possível, realizar uma análise do campo de estudos de insolvência e risco de crédito. Por meio deste estudo foi possível identificar os trabalhos, autores e periódicos mais relevantes dentro de tal campo de estudos, assim como a identificação de palavras-chave usadas nos estudos, técnicas que vêm sendo utilizadas na elaboração de estudos e tendências dentro do campo.

Diante disso, sugere-se que estudos futuros sobre o desenvolvimento do campo de estudos sejam realizados, para verificar se o campo continuará apresentando uma tendência de crescimento nos próximos anos. Ademais, sugere-se que estudos futuros de revisão voltados à técnicas computacionais como as de aprendizado de máquinas considerando as bases de dados *Scopus* e *Web of Science* sejam desenvolvidos. Por fim, sugere-se que os novos estudos empíricos do campo de estudos considerem os tópicos de tendência dos últimos anos, para a elaboração de modelos preditivos cada vez mais robustos, considerando também diferentes contextos, como o da pandemia de Covid-19.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALTMAN, E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, n. 4, p. 589–609, 1968.
- BAESENS, B. *et al.* Benchmarking state-of-the-art classification algorithms for credit scoring. **Journal of the Operational Research Society**, v. 54, n. 6, p. 627–635, jun. 2003.
- BALCAEN, S. OOGHE, H. 35 Years of Studies on Business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems. **The British Accounting Review**, n. 38, p. 63–93, 2006.
- BEAVER, W. H. Financial Ratios as Predictors of Failure. **Journal of Accounting Research**, v. 4, p. 71–111, 1966.
- BELLOVARY, J.; GIACOMINO, D.; AKERS, M. A Review of Bankruptcy Prediction Studies: 1930-Present. **Journal of Financial education**, v. 33, n. Winter, p. 1–42, 2007.
- BENOMAR, L. *et al.* Bibliometric Analysis of the Structure and Evolution of Research on Assisted Migration. **Current Forestry Reports**, v. 8, n. 2, p. 199–213, 2022.
- COBO, M. J. *et al.* An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the Fuzzy Sets Theory field. **Journal of Informetrics**, v. 5, n. 1, p. 146–166, 2011.
- DE LA HOZ-CORREA, A.; MUÑOZ-LEIVA, F.; BAKUCZ, M. Past themes and future trends in medical tourism research: A co-word analysis. **Tourism Management**, v. 65, p. 200–211, 2018.
- DEMYANYK, Y.; HASAN, I. Financial crises and bank failures: A review of prediction methods. **Omega**, v. 38, n. 5, p. 315–324, 2010.
- DIMITRAS, A. I. *et al.* Business failure prediction using rough sets. **European Journal of Operational Research**, v. 114, n. 2, p. 263–280, 1999.
- DIMITRAS, A. I.; ZANAKIS, S. H.; ZOPOUNIDIS, C. A survey of business failures with an emphasis on prediction methods and industrial applications. **European Journal of Operational Research**, v. 90, n. 3, p. 487–513, 1996.
- ECHCHAKOUI, S. Why and how to merge Scopus and Web of Science during bibliometric analysis: the case of sales force literature from 1912 to 2019. **Journal of Marketing Analytics**, v. 8, n. 3, p. 165–184, 2020.
- ECHCHAKOUI, S.; BARKA, N. Industry 4.0 and its impact in plastics industry: A literature review. **Journal of Industrial Information Integration**, v. 20, n. August 2019, p. 100172, 2020.
- FARIAS, F.; MARTÍNEZ, M. D. C.; MARTÍN-CERVANTES, P. A. Explanatory Factors of Business Failure: Literature Review and Global Trends. **Sustainability**, v. 13, n. 18, 2021.

- HAND, D. J.; HENLEY, W. E. Statistical Classification Methods in Consumer Credit Scoring: A Review. **Journal of Royal Statistical Society**, v. 160, n. 523–541, 1997.
- HUANG, C. L.; CHEN, M. C.; WANG, C. J. Credit scoring with a data mining approach based on support vector machines. **Expert Systems with Applications**, v. 33, n. 4, p. 847–856, 2007.
- KNANI, M.; ECHCHAKOUI, S.; LADHARI, R. Artificial intelligence in tourism and hospitality: Bibliometric analysis and research agenda. **International Journal of Hospitality Management**, v. 107, n. December 2021, p. 103317, 2022.
- KUMAR, P. R.; RAVI, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques - A review. **European Journal of Operational Research**, v. 180, n. 1, p. 1–28, jul. 2007.
- LESSMANN, S. et al. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. **European Journal of Operational Research**, v. 247, n. 1, p. 124–136, 2015
- LOMAX, K. S. Business failures - another example of the analysis of failure data. **Journal of the American Statistical Association**, v. 49, n. 268, p. 847–852, 1954.
- MAIA, S. C. et al. Mapping the literature on credit unions: a bibliometric investigation grounded in Scopus and Web of Science. **Scientometrics**, v. 120, n. 3, p. 929–960, 2019.
- MARBURG, T. F. A Study of Small Business Failure: Smith & Griggs of Waterbury. **Business History Review**, v. 28, n. 4, p. 366–384, 1954.
- MARX, W. et al. Detecting the historical roots of research fields by reference publication year spectroscopy (RPYS). **Journal of the Association for Information Science and Technology**, v. 65, n. 4, p. 751–764, 2014.
- MASTERS, G. N. A rasch model for partial credit scoring. **Psychometrika**, v. 47, n. 2, p. 149–174, 1982.
- MIN, J. H.; LEE, Y.-C. Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. **Expert Systems with Applications**, v. 28, n. 4, p. 603–614, 2005.
- MOHER, D. et al. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. **International Journal of Surgery**, v. 8, n. 5, p. 336–341, 2010.
- MURCIA, F. C. DE S.; MURCIA, F. D.-R.; BORBA, J. A. Rating de crédito corporativo: revisão da literatura e oportunidades para pesquisa no cenário brasileiro. **Revista de Economia e Administração**, v. 13, n. 1, p. 54–96, 2014.
- OHLSON, J. A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. **Journal of Accounting Research**, v. 18, n. 1, p. 109–131, 1980.
- PINTO, M.; PULGARÍN, A.; ESCALONA, M. I. Viewing information literacy concepts: A comparison of two branches of knowledge. **Scientometrics**, v. 98, n. 3, p. 2311–2329, 2014.

PRADO, J. *et al.* Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). **Scientometrics, Budapeste**, v. 106, p. 1007–1029, 2016.

SHI, Y.; LI, X. An overview of bankruptcy prediction models for corporate firms: A systematic literature review. **Intangible Capital**, v. 15, n. 2, SI, p. 114–127, 2019a.

SHI, Y.; LI, X. A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. **Heliyon**, v. 5, n. 12, 2019b.

SHEPHERD, D. A. Learning from Business Failure: Propositions of Grief Recovery for the Self-Employed. **The Academy of Management Review**, v. 28, n. 2, p. 318–328, 21 fev. 2003.

SHUMWAY, T. Forecasting bankruptcy more accurately: A simple hazard model. **Journal of Business**, v. 74, n. 1, p. 101–124, jan. 2001.

UBARHANDE, P.; CHANDANI, A. Elements of Credit Rating: A Hybrid Review and Future Research Agenda. **Cogent Business and Management**, v. 8, n. 1, 2021.

VAN ECK, N. J.; WALTMAN, L. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, v. 84, n. 2, p. 523–538, 2010.

VANTI, N. A. P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, v. 31, n. 2, p. 152–162, 2002.

WEST, D. Neural network credit scoring models. **Computers & Operations Research**, v. 27, p. 1131–1152, 2000.

ZMIJEWSKI, M. E. Methodological Issues Related to the Estimation of Financial Distress Prediction Models. **Journal of Accounting Research**, v. 22, p. 59–82, 1984.

ZUPIC, I.; ČATER, T. Bibliometric Methods in Management and Organization. **Organizational Research Methods**, v. 18, n. 3, p. 429–472, 2015.

ARTIGO 2 – RISCO DE INSOLVÊNCIA: UMA ANÁLISE COMPARATIVA DE DIFERENTES TÉCNICAS DE PREVISÃO PARA O MOMENTO DE PANDEMIA

RESUMO

O objetivo desse artigo foi comparar técnicas estatísticas e de aprendizado de máquina das três classificações principais: caixa branca, caixa cinza e caixa preta, considerando o período de pandemia como amostra de teste. Para o desenvolvimento do estudo, foram usadas as técnicas de análise discriminante, regressão logística, k-vizinhos mais próximos, árvore de decisão, redes neurais artificiais e máquinas de vetores de suporte. Para avaliar as técnicas, a acurácia e a curva ROC de cada um dos modelos foram analisadas. Como resultado, a técnica com melhor percentual de acurácia foi a árvore de decisão, seguida pela de floresta aleatória, regressão logística e análise discriminante. Por fim, estiveram a rede neural artificial (RNA), K-vizinhos mais próximos (KNN) e máquina de vetores de suporte (SVM). Tal resultado mostrou que técnicas de caixa branca, caixa cinza e técnicas estatísticas apresentaram desempenhos melhores de acurácia em relação às técnicas de caixa preta (RNA e SVM). Sendo assim, considerando as empresas que entraram em situação de insolvência no período de pandemia como amostra de teste, os resultados apontaram para a superioridade de uma técnica de caixa branca, seguida de uma de caixa cinza e das técnicas estatísticas tradicionais. Diante da existência de uma ampla variedade de técnicas para a elaboração dos modelos preditivos, torna-se relevante ter conhecimento sobre qual fornece maiores resultados de previsão para determinadas amostras. Sendo assim, é importante tentar identificar entre as técnicas existentes qual fornece melhores resultados de classificação diante de contextos (como o da pandemia) e empresas específicos.

Palavras-chave: Aprendizado de máquinas. Previsão de falência. COVID-19.

ABSTRACT

The aim of this article was to compare statistical and machine learning techniques of the three main classifications: white box, gray box and black box, considering the pandemic period as a test sample. For the development of the study, the techniques of discriminant analysis, logistic regression, k-nearest neighbors, decision tree, artificial neural networks and support vector machines were used. To evaluate the techniques, the accuracy and the ROC curve of each of the models were analyzed. As a result, the technique with the best percentage of accuracy was the decision tree, followed by the random forest, logistic regression and discriminant analysis. Finally, there were the artificial neural network (ANN), K-nearest neighbors (KNN) and support vector machine (SVM). This result showed that white box, gray box and statistical techniques presented better performance in terms of accuracy compared to black box techniques (ANN and SVM). Therefore, considering companies that went bankrupt during the pandemic period as a test sample, the results pointed to the superiority of a white box technique, followed by a gray box and traditional statistical techniques. Given the existence of a wide variety of techniques for the development of predictive models, it becomes relevant to know which one provides the best forecast results for certain samples. Therefore, it is important to try to identify, among the existing techniques, which provides better classification results in contexts (such as the pandemic) and specific companies.

Keywords: Machine learning. Bankruptcy prediction. COVID-19.

1 INTRODUÇÃO

A previsão de insolvência se constitui em um campo de estudos que tem atraído um volume crescente de pesquisas. De acordo com Battiston *et al.* (2007) a falência de uma empresa pode causar um efeito dominó (sistêmico), uma vez que existem relações em cadeia com outras empresas. Ou seja, a falência de uma empresa impossibilita o pagamento de seus fornecedores, o que compromete a solvência desses fornecedores e de toda a cadeia.

A previsão de insolvência será, portanto, relevante para todos os agentes envolvidos com uma empresa que possa se tornar insolvente (BATTISTON *et al.*, 2007; HORTA; ALVES; CARVALHO, 2014), sendo eles tomadores de decisão, investidores, fornecedores, credores, funcionários, ou seja, para todos os *stakeholders* da empresa (PRADO *et al.*, 2016).

Devido à relevância do tema, diversos estudos de previsão têm sido desenvolvidos utilizando diferentes técnicas (PRADO *et al.*, 2016). Os estudos iniciais utilizavam principalmente técnicas estatísticas como a análise discriminante, usada no trabalho seminal de Altman (1968) e regressão logística, introduzida por Olshon (1980). Nos últimos anos, os esforços para a criação de modelos mais robustos têm sido crescentes, com o uso do aprendizado de máquina, com técnicas como redes neurais artificiais (RNAs) e máquinas de vetores de suporte (*support vector machines* – SVMs).

De acordo com Obermann e Waack (2015), o aprendizado de máquina é entendido como um método algorítmico que tem como objetivo a generalização. Por meio da troca de amostras de aprendizado e do novo cálculo dos classificadores, variações como mudanças econômicas ou as necessidades de diferentes setores podem ser identificadas pelo programa sem a modificação de sua lógica. As decisões tomadas pelas técnicas de aprendizado de máquina têm como característica a objetividade, que tem os dados em análise como base.

Existem, no entanto, diferentes técnicas de aprendizado de máquina que são classificadas em três categorias principais, que são técnicas de caixa branca, caixa cinza e caixa preta (OBERMANN; WAACK, 2015; SJÖBERG *et al.*, 1995). A ideia dessa classificação tem como base a interpretabilidade por trás de cada técnica. Sendo assim, geralmente assume-se que as técnicas de caixa branca são as mais interpretáveis e menos robustas, enquanto as de caixa cinza são intermediárias e as de caixa preta são menos interpretáveis e mais robustas (OBERMANN; WAACK, 2015).

Face a tais diferenças, e mesmo diante da ideia de superioridade das técnicas de caixa preta, a comparação de técnicas de previsão também se torna relevante. Isso se deve pelo fato de que existem evidências de estudos na área de finanças que não confirmam a superioridade

de modelos menos interpretáveis quando comparados à modelos interpretáveis (OBERMANN; WAACK, 2015). Sendo assim, por meio da comparação entre as técnicas se torna possível identificar qual fornece resultados mais satisfatórios para uma determinada amostra.

Apesar de uma previsão de insolvência assertiva diante a eventos não previstos ser de difícil mensuração (KITOWSKI; KOWAL-PAWUL; LICHOTA, 2022), a previsão pode ser relevante para a antecipação de problemas de insolvência. Isso porque em momentos de recessão, como os causados pela pandemia de COVID-19, tal antecipação é considerada determinante para que as empresas consigam evitar a falência, aderindo, por exemplo políticas públicas governamentais como as desenvolvidas especialmente para o momento de pandemia e também outras estratégias. Por isso, é relevante analisar também se entre as técnicas de previsão, existem técnicas que geram resultados robustos mesmo diante de situações imprevisíveis, como a pandemia de COVID-19.

Para contribuir com a literatura, o presente estudo teve como objetivo comparar técnicas estatísticas (análise discriminante, regressão logística) e de aprendizado de máquina das três classificações principais: Caixa branca (k-vizinhos mais próximos e árvore de decisão), caixa cinza (floresta aleatória) e caixa preta (redes neurais e máquina de vetores de suporte), considerando o período de pandemia como amostra de teste para avaliar os modelos de previsão.

As técnicas de aprendizado de máquina foram escolhidas considerando a classificação baseada na interpretabilidade, com o intuito de analisar se as técnicas de caixa preta realmente fornecem resultados mais satisfatórios, quando comparadas com técnicas estatísticas, de caixa branca ou caixa cinza. Ademais, optou-se por essas técnicas porque elas já foram usadas em estudos de previsão de insolvência anteriores (BARBOZA; KIMURA; ALTMAN, 2017; OBERMANN; WAACK, 2015; PINTO, 2021; SANTOS, 2021).

Para o alcance do objetivo foram desenvolvidos os seguintes passos: Definir os indicadores relevantes para a elaboração dos modelos preditivos; comparar as técnicas análise discriminante (LDA), regressão logística (LR), k-vizinhos mais próximos (KNN), árvore de decisão (DT), floresta aleatória (RF), redes neurais artificiais (RNA) e máquina de vetores de suporte (SVM); analisar a robustez das técnicas usadas, verificando os resultados de previsão de insolvência para o período de pandemia de COVID-19.

Diante da existência de uma ampla variedade de técnicas para a elaboração dos modelos preditivos, torna-se relevante ter conhecimento sobre qual fornece resultados de previsão mais robustos para determinadas amostras. Isso se faz relevante devido às especificidades das empresas como tamanho, setor de atuação e forma legal. Sendo assim, é importante identificar

entre as técnicas existentes qual fornece melhores resultados de classificação diante de contextos e empresas específicos, que possuam características em comum.

2 REFERENCIAL TEÓRICO

Trabalhos sobre previsão de insolvência têm sido desenvolvidos ao longo do tempo. Um dos primeiros trabalhos sobre a temática de previsão de falência foi o de Beaver (1966). O estudo usou a análise univariada, com uma amostra composta por 158 empresas falidas e não falidas. Outro trabalho que está entre os seminais e que é o mais influente sobre a temática, de acordo com Prado *et al.* (2016), é o de Altman (1968). Esse estudo foi o primeiro a usar uma técnica multivariada (análise discriminante) para a elaboração do modelo preditivo.

A partir do estudo, foi desenvolvido o índice geral Z, composto por cinco indicadores econômico-financeiros. O modelo alcançou um percentual de acerto de 95% considerando dados com um ano de antecedência do evento de insolvência, percentual que de acordo com o autor diminuía à medida que os dados fossem de anos mais distantes do evento de insolvência.

O trabalho de Olshon (1980) também foi relevante para a temática. Este estudo foi o primeiro a usar um modelo *logit* para a elaboração do modelo preditivo. Nele o autor considerou uma amostra formada por 105 empresas insolventes e 2.058 empresas solventes. Em relação ao nível de acertos do modelo, o autor mencionou que o ponto de corte em que os erros foram minimizados foi de 0,038. Nele, foram classificadas incorretamente 17,4% das empresas solventes e 12,4% das insolventes. Isso indica que considerando uma população de empresas com o mesmo número de empresas solventes e insolventes, a taxa de erro esperada seria de 14,9%.

O trabalho de Odom e Sharda (1990) foi o primeiro a usar a técnica de redes neurais artificiais para a elaboração de um modelo de previsão de insolvência. Face à habilidade preditiva das redes neurais artificiais, os autores elaboraram um modelo usando dados financeiros de 129 empresas (65 insolventes e 64 solventes), considerando o período de 1975 a 1982.

De maneira complementar, os autores usaram a técnica de análise discriminante (tradicionalmente usada em modelos de previsão) para comparar com os resultados das redes neurais. Como resultado, os autores verificaram que o uso das redes neurais artificiais para a previsão de insolvência era promissor. Tal constatação foi feita porque no estudo os resultados da rede neural artificial superaram os resultados da análise discriminante nas amostras de teste e de treinamento.

No trabalho de Hillegeist *et al.* (2004), os autores compararam os modelos de Altman (1968) – *Z-score* – e de Ohlson (1980) – *O-score* – com uma medida de probabilidade de falência (PB) derivada do modelo de *Black-Scholes-Merton* (BSM-Prob), de precificação de

opções. Para a análise, os autores usaram uma amostra com 78.100 observações do período de 1980 a 2000 e a metodologia de taxa de risco discreta. A partir dos resultados, os autores concluíram que o BSM-Prob apresentou resultados mais robustos que os modelos baseados em indicadores contábeis (*Z-score* e *O-score*).

No trabalho de Wilson e Sharda (1994), os autores compararam as técnicas de redes neurais artificiais e análise discriminante. Para a comparação, os autores optaram por usar os mesmos indicadores econômico-financeiros usados no modelo de Altman (1968). A amostra do estudo foi composta por 129 empresas, sendo 65 delas insolventes e 64 solventes e os dados foram correspondentes aos de aproximadamente um ano antes da situação de insolvência. A partir do estudo os autores concluíram que a técnica de redes neurais artificiais apresentou resultados mais robustos para a previsão de empresas solventes e insolventes, comparada à análise discriminante.

Outro trabalho que usou uma técnica diferenciada comparando-a com técnicas até então usadas foi o de Dimitras *et al.* (1999). Nele os autores usaram a técnica chamada *rough set approach*. O objetivo do estudo foi analisar se a *rough set approach* apresentaria resultados robustos para a previsão de insolvência, e comparando-a com as técnicas de análise discriminante e regressão logística (*logit*).

Para a análise, 40 empresas gregas e insolventes, e 40 empresas gregas solventes foram consideradas. Como resultado, os autores verificaram que a técnica de *rough set approach* apresentou resultados melhores que a análise discriminante e a análise *logit*. Apesar disso, a superioridade do modelo em relação à análise discriminante foi maior que a análise *logit*.

No trabalho de West (2000), o autor comparou a precisão de cinco arquiteturas de redes neurais artificiais com métodos mais tradicionais para a pontuação de crédito. Os métodos tradicionais analisados no estudo foram análise discriminante, regressão logística, k vizinho mais próximo, estimativa de densidade de Kernel e árvore de decisão. Para a análise, foram considerados dados de crédito alemães e australianos.

Diferentemente dos trabalhos citados anteriormente, este estudo não analisou empresas, mas sim pessoas. Da amostra de alemães (1.000 no total), 700 eram indivíduos para os quais o crédito poderia ser concedido e 300 que não. Da amostra de australianos (690 no total), para 307 indivíduos o crédito poderia ser concedido e para 383 não.

A partir do estudo, o autor concluiu que as redes neurais apresentaram melhores resultados quanto à precisão na pontuação de crédito, necessitando, porém de habilidades em sua modelagem para o alcance desses resultados. Ademais, ao considerar as técnicas tradicionais, a que apresentou os melhores resultados foi a de regressão logística, sendo,

portanto, uma alternativa às redes neurais. Outro fato destacado foi que os modelos não-paramétricos k-vizinhos mais próximos e densidade de Kernel não produziram resultados satisfatórios.

O trabalho de Baesens *et al.* (2003) analisou a classificação de crédito usando dados das principais instituições financeiras da Bélgica, Holanda, Luxemburgo e Reino Unido. No estudo os autores também realizaram uma comparação de diversas técnicas, que foram regressão logística, análise discriminante linear e quadrática, programação linear (mínimos quadrados), máquinas de vetor de suporte (SVMs), redes neurais (árvore aumentada) naive Bayes e classificadores de vizinhos mais próximos.

A partir do estudo, foi verificado que os diferentes tipos de técnicas apresentaram resultados diferentes estatisticamente. Os melhores resultados foram os das técnicas máquinas de vetor de suporte de mínimos quadrados (LS-SVMs) e redes neurais. No entanto, técnicas tradicionais mais simples também apresentaram resultados bastante satisfatórios, sendo elas regressão logística e análise discriminante linear.

No estudo de Huang *et al.* (2004), os autores compararam as técnicas de máquinas de vetor de suporte (SVMs) e redes neurais de retropropagação para a previsão de classificação de crédito. Para a análise, os autores usaram dados de empresas dos mercados dos Estados Unidos e Taiwan, que foram classificadas como AA, A, BBB, BB ou B. A partir dos resultados, os autores verificaram que a técnica de máquinas de vetor de suporte apresentou resultados satisfatórios, com precisão comparável à técnica de redes neurais de retropropagação.

O estudo de Shin, Lee e Kim (2005) apresentou uma investigação a cerca da eficiência da técnica de máquinas de vetor de suporte para a previsão de insolvência. Assim como no estudo de Huang *et al.* (2004), os autores compararam a precisão das técnicas de máquinas de vetor de suporte e redes neurais de retropropagação. Os autores usaram dados de 2.320 empresas coreanas, sendo que 1.160 empresas apresentaram pedido de falência e 1.160 não pediram. O período analisado foi de 1996 a 1999. Como resultado, os autores concluíram que a medida que a amostra de treinamento ficou menor, a técnica de máquinas de vetor de suporte apresentou resultados mais robustos que a de redes neurais de retropropagação.

Em estudo mais recente, Obermann e Waack (2015) realizaram uma comparação entre técnicas de aprendizado de máquina. Segundo os autores, existem três tipos principais de técnicas de aprendizado de máquina (caixa branca, caixa cinza e caixa preta). Segundo eles, é comum que seja assumido que as técnicas de caixa branca sejam mais interpretáveis e menos robustas, enquanto as de caixa preta sejam menos interpretáveis e mais robustas.

Usando um banco de dados com 5.152 empresas, os autores verificaram que, diferentemente do geralmente proposto, as técnicas de caixa branca não apresentaram resultados inferiores às de caixa preta. Tal resultado mostrou que métodos interpretáveis são apropriados para a elaboração de modelos preditivos.

Outro estudo que comparou técnicas de previsão foi o de Barboza, Kimura e Altman (2017). No estudo, os autores compararam técnicas tradicionais (análise discriminante, regressão logística e redes neurais) com técnicas de aprendizado de máquina (SVM, *bagging*, *boosting*, e floresta aleatória). Para a análise, os autores usaram dados de empresas norte-americanas no período de 1985 a 2013. A partir do estudo, os autores concluíram que as técnicas de aprendizado de máquina apresentaram resultados mais robustos (cerca de 10% mais precisas que as técnicas tradicionais). Ademais, dentre as técnicas de aprendizado de máquina analisadas, as que apresentaram resultados superiores foram as de *bagging*, *boosting* e floresta aleatória.

Considerando o cenário brasileiro, os estudos de Pinto (2021) e Santos (2021) desenvolveram modelos de previsão considerando diferentes técnicas de previsão. Especificamente, no trabalho de Pinto (2021) o autor desenvolveu modelos de previsão de insolvência considerando técnicas tradicionais e de aprendizado de máquina, para a comparação entre elas.

No estudo, as técnicas analisadas foram *naive bayes*, SVM Linear, SVM *polynomial*, SVM radial, regressão logística, redes neurais, floresta aleatória e *XGBoost*. Usando uma amostra composta por empresas do setor de atacado, o autor identificou que a técnica com melhor poder preditivo para a amostra analisada foi o *XGBoost*, que apresentou um poder preditivo de 98,38%.

No trabalho de Santos (2021), o autor teve como objetivo comparar o poder preditivo de técnicas de aprendizado de máquina para o risco de insolvência, considerando dados de empresas brasileiras de capital aberto. O autor analisou as técnicas de floresta aleatória, *naive bayes*, *logit*, K-NN, SVM (linear, polinomial e de base radial), *bagging* e *boosting*. Por meio do estudo, o autor verificou que a técnica de floresta aleatória foi a que apresentou maior assertividade, com um total de 95,72%.

2.1 Técnicas estatísticas para previsão

Entre as técnicas estatísticas usadas para a previsão de insolvência, as técnicas de análise multivariada se destacam. De acordo com Hair *et al.* (2006), a análise multivariada é entendida

como uma análise em que diversas variáveis são usadas em uma ou mais relações. Ou seja, a análise multivariada envolve todas as técnicas estatísticas em que diversas medidas (variáveis) sobre objetos ou indivíduos são analisadas ao mesmo tempo. Dessa maneira, qualquer análise estatística envolvendo a análise de mais que duas variáveis ao mesmo tempo é considerada como uma análise multivariada.

No campo da previsão de insolvência, Altman (1968) introduziu a análise multivariada, usando em seu estudo a técnica de análise discriminante. Posteriormente foram introduzidas também outras técnicas estatísticas, como a regressão logística (OHLSON, 1980) e também técnicas de aprendizado de máquina (ODOM; SHARDA, 1990). Apesar da introdução de novas técnicas, de acordo com Prado *et al.* (2016), a análise discriminante e a regressão logística continuaram sendo usadas.

2.1.1 Análise discriminante

De acordo com Corrar, Paulo e Dias Filho (2007), a análise discriminante é entendida como uma técnica estatística a partir da qual é possível a identificação de variáveis que diferenciam os grupos. Segundo os mesmos autores, a análise discriminante também possibilita a identificação das variáveis necessárias para a classificação. Ou seja, a análise discriminante determina as variáveis relevantes e que são capazes de discriminar os grupos.

Hair *et al.* (2006) afirmam que a análise discriminante, assim como a regressão logística, é uma técnica apropriada quando a variável dependente é nominal ou não-métrica e as variáveis explicativas são métricas. A classificação de crédito é um exemplo, pois as variáveis explicativas geralmente são indicadores econômico – financeiros e a variável dependente é a classificação da empresa como boa (solvente) ou ruim (insolvente). A função discriminante segundo Hair *et al.* (2006) é dada pela Equação 1:

$$Z_{jk} = a + W_1X_{1k} + W_2X_{2k} + \dots + W_nX_{n1k} \quad (1)$$

Em que:

Z_{jk} = score Z discriminante da função discriminante j para o objeto k ;

a = intercepto;

W_1 = peso discriminante para a variável independente i ;

X_{1k} = variável explicativa i para o objeto k .

Conforme pode ser visto na Equação 1, Hair *et al.* (2006) afirmam que na análise discriminante são estabelecidos pesos para as variáveis explicativas, os quais objetivam a maximização das diferenças entre os grupos. Ainda segundo Hair *et al.* (2006), o valor dado pela função discriminante para cada objeto que está sendo analisado é obtido pela soma dos produtos de cada peso com sua respectiva variável explicativa. Ou seja, o escore discriminante é o resultado das somas das multiplicações entre W_n e X_{nk} .

A partir da introdução da técnica de análise discriminante por Altman (1968), diversos estudos utilizaram essa técnica para a elaboração de modelos de previsão. Entre estes estudos podem-se citar os de Kanitz (1974), Matias (1978), Silva (1982), Sanvicente e Minardi (1998) e Stüpp (2015).

2.1.2 Regressão logística (*logit*)

A regressão logística, também conhecida como *logit*, foi introduzida na pesquisa de previsão de insolvência por Ohlson (1980). De acordo com Hair *et al.* (2006), a regressão logística é uma forma específica de regressão para a previsão e explicação de uma variável não-métrica binária (que possui dois grupos).

Nessa técnica, a função estatística conta com uma relação multivariada, parecida com a de uma regressão comum, relação esta que apresenta coeficientes que indicam o impacto de cada variável explicativa. Corroborando o mencionado, Prado (2016) afirma que a regressão logística pode ser entendida como uma modelagem matemática multivariada que possibilita a identificação da probabilidade de um evento. A função que representa a regressão logística, está representada pela Equação 2.

$$P(Y = 1) = \frac{1}{1 + e^{-(x)}} \quad (2)$$

Em que:

$P(Y=1)$ = probabilidade relativa que um determinado evento ocorra;

$e \cong 2,718$;

$(x) = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni}$

Em que: β_0 = uma constante; β_n = coeficientes estimados por meio dos valores de X_{ni} .

A regra de classificação usada para diferenciar dois grupos na regressão logística é dada por: $P(Y=1) > 0,5$, então $Y=1$; se $P(Y=1) < 0,5$, então $Y=0$. A probabilidade (P) está situada entre zero e um (PRADO, 2019). Assim o escore produzido pela regressão logística pode ser entendido como a probabilidade de ocorrência do evento que está sendo analisado (por exemplo, para identificar a probabilidade de uma empresa se tornar insolvente, em que: insolvente=1; solvente=0).

Após o trabalho de Olshon (1980) outros autores também desenvolveram estudos que usavam a regressão logística para a elaboração de modelos de previsão. Entre eles estiveram Akiama (2008), Stüpp (2015), Barboza, Kimura e Altman (2017) e Santos (2021).

2.2 Técnicas de aprendizado de máquina

De acordo com Izbick e Santos (2020), o nascimento do aprendizado de máquina se deu nos anos 60. Segundo os autores, esse era um campo da inteligência artificial em que padrões eram aprendidos por meio de dados. A partir da década de 90 houve uma expansão das técnicas de aprendizado de máquina e suas aplicações passaram a ter pontos em comum com as técnicas estatísticas. Sendo assim, passou a ser comum a aplicação de métodos de aprendizado de máquina em estatística.

Teles *et al.* (2021) afirmam que o aprendizado de máquina é entendido como um campo da inteligência artificial em que há o desenvolvimento de métodos, estratégias e algoritmos. Este campo, segundo os autores, possui uma forte ligação com a estatística, existindo também diferentes técnicas que podem ser aplicadas. No aprendizado de máquina, o sistema computacional aprende com os dados e executa atividades tendo como base os dados da amostra juntamente com a aplicação de testes nos novos dados (TELES *et al.*, 2021).

De maneira complementar, Barbosa (2009) afirmou que o principal objetivo do campo de aprendizado de máquina consiste no estudo e desenvolvimento de métodos computacionais para a construção de sistemas capazes de aprender automaticamente. Ainda segundo o autor, a maximização da capacidade de generalização é um grande desafio para tais algoritmos. Obermann e Waack (2015) complementam que existe um alto nível de objetividade no aprendizado de máquina, que segundo os autores é um método algorítmico que tem como objetivo a generalização, por meio de hipóteses derivadas de dados de treinamento.

As técnicas de aprendizado de máquina são classificadas em três categorias principais, que são as técnicas de caixa branca, caixa cinza e caixa preta (OBERMANN; WAACK, 2015; SJÖBERG *et al.*, 1995). Essa classificação se baseia na interpretabilidade de cada método, em

que os modelos de caixa branca são assumidos como mais interpretáveis (e menos robustos) e os de caixa preta como menos interpretáveis (e mais robustos).

A interpretabilidade é entendida como o grau de entendimento que os seres humanos conseguem obter em relação à decisão do modelo, ou seja, quanto o ser humano consegue entender do raciocínio por trás do resultado que determinado modelo fornece (MILLER, 2019; KANAPRAM *et al.*, 2022). Sendo assim, a principal ideia assumida era de que as técnicas menos interpretáveis pelos seres humanos apresentariam resultados superiores às técnicas mais interpretáveis.

Nas técnicas de aprendizado de máquina geralmente existe uma troca entre interpretabilidade e precisão (OBERMANN; WAACK, 2015). Isso significa que modelos precisos perdem precisão ao se tornar mais interpretáveis, e modelos interpretáveis ganham precisão ao tornarem menos interpretáveis. Apesar da existência desse pressuposto, Obermann e Waack (2015) mencionaram que na área de finanças existem evidências que mostram que não há, necessariamente, superioridade de modelos não interpretáveis em relação a modelos interpretáveis.

Devido à existência de diferentes técnicas de aprendizado de máquina, se faz relevante analisar quais fornecem os melhores resultados para amostras específicas. Ademais, a comparação das técnicas de aprendizado de máquina e de técnicas consideradas tradicionais também pode oferecer importantes perspectivas relacionadas à previsão de insolvência. Para isso, nesse estudo serão usadas técnicas de aprendizado de máquina de caixa branca (árvore de decisão e k-vizinhos mais próximos), caixa cinza (floresta aleatória) e caixa preta (redes neurais artificiais e máquinas de vetor de suporte), que estão detalhadas adiante.

2.2.1 K-vizinhos mais próximos (K-nearest neighbors – KNN)

O KNN é uma técnica que, de acordo com Müller e Guido (2017), trata-se do mais simples algoritmo de aprendizado de máquina. Nessa técnica, considerada como de caixa branca (WEŁESZCZUK; KOSIŃSKA-SELBI; CHOLEWIŃSKA, 2022), o algoritmo encontra seus “vizinhos mais próximos”, para a fazer a previsão de um novo ponto de dados. Ou seja, nesse caso, a classificação do ponto de dados se dá por meio da classificação que prevalece entre o número de vizinhos mais próximos que compuseram a amostra de treinamento. Essa técnica, segundo Izbick e Santos (2020), pode ser representada matematicamente pela Equação 1.

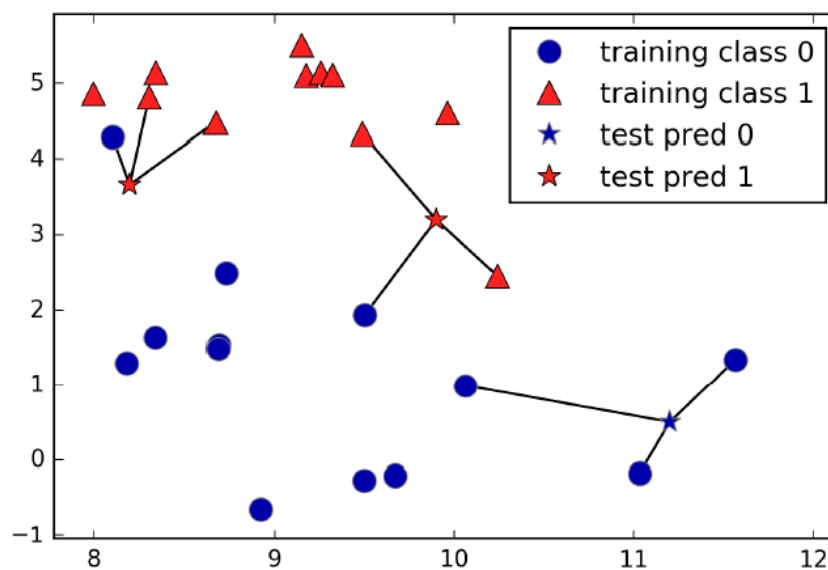
$$g(x) = \frac{1}{k} \sum_{i \in N_x} y_i \quad (1)$$

Em que:

N_x = conjunto das k observações mais próximas de x.

A Figura 18 mostra a representação do funcionamento da técnica KNN.

Figura 18 — Funcionamento da técnica KNN.



Nota: *Training class 0* = classificação 0 na amostra de treinamento.

Training class 1: classificação 1 na amostra de treinamento.

Test pred 0: previsão 0 na amostra de teste.

Test pred 1: previsão 1 na amostra de teste.

Fonte: Müller e Guido (2017, p. 36).

Com base na Figura 18, percebe-se que a classificação final se dará por meio da classificação predominante entre os vizinhos que compuseram a amostra de treinamento. O número de vizinhos mais próximos considerados será um fator determinante. Na Figura 18, o número de vizinhos mais próximos considerado foi de três, assim, a classificação seguiu a classificação que predominou entre os três vizinhos mais próximos.

2.2.2 Árvore de decisão (*Decision tree - DT*)

O algoritmo de árvore de decisão foi desenvolvido por Quinlan (1986). Segundo o autor, essa metodologia era relativamente simples, mas capaz de resolver problemas considerados

como difíceis na prática. De acordo com Obermann e Waack (2015), o algoritmo de árvore de decisão pode ser considerado como uma técnica de aprendizado de máquina de caixa branca. Isso indica que essa técnica possui boa interpretabilidade.

De acordo com Quinlan (1986), a estratégia usada nessa técnica é o aprendizado não incremental, obtido por meio de exemplos. Segundo o autor, um conjunto de casos a serem classificados são apresentados ao sistema, que desenvolve uma árvore de decisão. De maneira resumida, pode ser entendido que as árvores de decisão desenvolvidas pelo sistema são responsáveis pela classificação com base nos dados disponibilizados. Essas árvores tem a raiz como ponto de partida, terminando nas folhas. De acordo com Izbick e Santos (2020), a árvore de decisão pode ser representada pela Equação 3.

$$g(x) = \text{moda}\{y_i: x_i \in R_k\} \quad (3)$$

Em que:

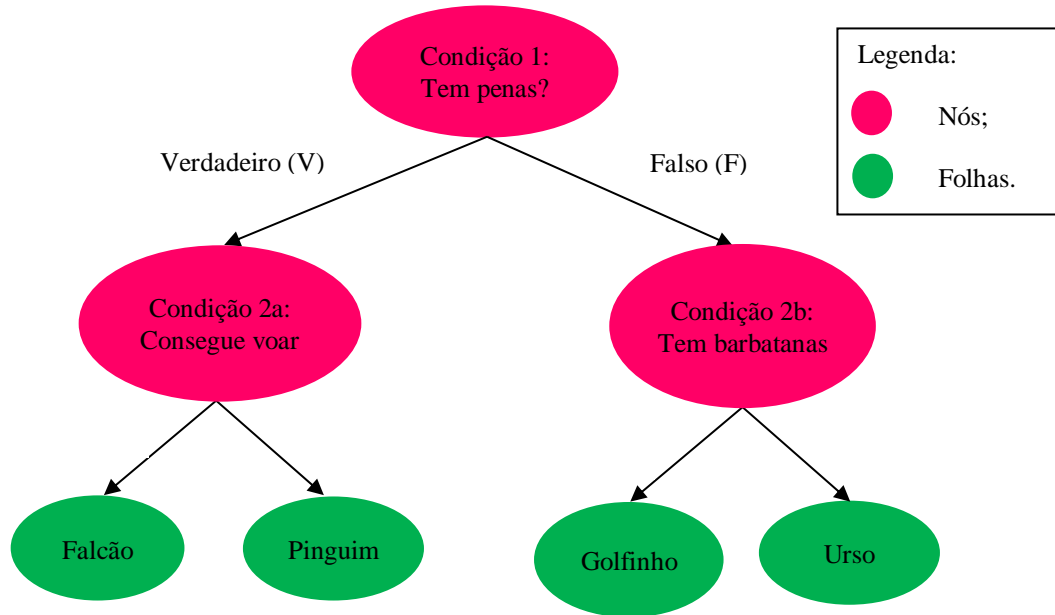
y_i = resposta Y de uma observação;

x_i = resposta X de uma observação;

R_k = região.

De acordo com Mueller e Massaron (2021), o algoritmo de árvore de decisão possui uma estrutura com ramificações que ilustram os resultados da decisão. Nessa técnica, os resultados de uma decisão podem ser mapeados e cada nó de uma árvore de decisão representa um resultado possível, sendo que em cada nó são atribuídas probabilidades de que o resultado ocorra. A estrutura de uma árvore de decisão pode ser exemplificada pela Figura 19.

Figura 19 — Árvore de decisão para a classificação de animais.



Fonte: Adaptado de Mueller e Guido (2019, p. 71) e Izbick e Santos (2020, p. 77).

De acordo com Izbick e Santos (2020), o funcionamento de uma árvore para a previsão de uma observação termina quando uma folha é atingida. Ou seja, a técnica se inicia com uma condição em seu topo (pode ser considerado como a raiz – primeiro nó). Essa condição será analisada e, caso seja verdadeira, chega-se à esquerda, caso não seja, a análise continua e uma nova condição (segundo nó, à direita, que pode ser considerado como galho) é verificada, conforme mostrado na Figura 19. Esse raciocínio é continuado até que se chegue a uma classificação final, chamada de folha.

Ainda segundo Izbick e Santos (2020), para a construção de uma árvore, uma etapa de grande importância é a de poda, que evita o super ajuste. Esse super ajuste atrapalha a eficiência preditiva em observações novas, por isso sua variância deve ser minimizada. A poda funciona como um meio para tornar a árvore menos complexa e menor, diminuindo também a variância do super ajuste.

Face à atenção dada às técnicas de aprendizado de máquina, alguns estudos que utilizaram a técnica de árvore de decisão para a elaboração de modelos de previsão de insolvência são encontrados. Entre eles estão os estudos de Aoki e Hosonuma (2004), Cho, Hong e Ha (2010), Chen (2012), Obermann e Waack (2015), Gavurova et al. (2017) e Korol (2019).

2.2.3 Floresta aleatória (*Random Forest* - RF)

A técnica de floresta aleatória (RF – *random forest*) foi introduzida por Breiman (2001). De acordo com o autor, as florestas aleatórias são formadas por preditores de árvores. As árvores dependem, individualmente, de valores de um vetor aleatório independente, que possui a mesma distribuição para as árvores da floresta.

De maneira complementar, Obermann e Waack (2015) mencionam que as florestas aleatórias são compostas por um conjunto de árvores de decisão. Diferentemente do algoritmo de árvore de decisão, no entanto, na floresta aleatória não há poda e o resultado se baseia no resultado da classificação das árvores em conjunto, em que prevalece a opção “mais votada”.

Corroborando o mencionado, Izbick e Santos (2020) afirmam que a floresta aleatória é um método em que diversas árvores são combinadas para realizar a predição de um problema. Isso significa que nesse algoritmo várias árvores diferentes são criadas. A partir da combinação dos resultados individuais de cada árvore, a predição é realizada, sendo, portanto, mais robusto que uma única árvore. A representação da técnica de floresta aleatória pode ser representada pela Equação 4, de acordo com Izbick e Santos (2020).

$$g(\mathbf{x}) = \text{moda}\{g^b(x), b = 1, \dots, B\} \quad (4)$$

Em que “uma observação com covariáveis \mathbf{x} é classificada por cada árvore construída e, posteriormente, a predição é dada pela categoria predita com maior frequência” (IZBICK; SANTOS, 2020, p. 159).

De acordo com Obermann e Waack (2015), apesar de ser considerada como um método mais robusto que árvores de decisão, a floresta aleatória apresenta alguns problemas que fazem com que ela se torne menos interpretável. O principal deles é o uso de várias árvores que possuem características diferentes. Outro problema consiste no tamanho de cada árvore, pois uma vez que não há poda, essas árvores se tornam maiores. No entanto, mesmo com problemas como os mencionados, os autores afirmam que a técnica de floresta aleatória possui interpretabilidade. Sendo assim, essa técnica pode ser considerada como uma intermediária entre as técnicas de caixa branca e caixa preta, sendo, portanto, uma técnica de caixa cinza.

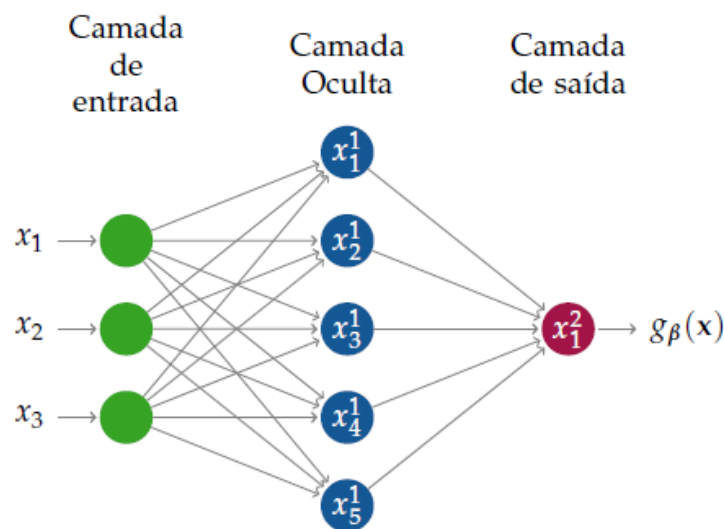
Entre os estudos que utilizaram a técnica de floresta aleatória, podem ser citados os estudos de Obermann e Waack (2015), Barboza, Kimura e Altman (2017), Pinto (2021) e Santos (2021). Nesses estudos a comparação entre técnicas foi realizada, no estudo Barboza, Kimura e Altman (2017), a técnica de floresta aleatória esteve entre as três com melhores

assertividade, e no trabalho de Santos (2021), ela foi a técnica com maior assertividade entre as comparadas.

2.2.4 Redes neurais artificiais (RNAs)

As redes neurais artificiais são uma técnica de inteligência artificial não algorítmica que funcionam de forma análoga ao funcionamento de neurônio humano (BRAGA; LUDERMIR; CARVALHO, 2000). De maneira complementar, Bishop (1995) afirmou que as redes neurais são um método computacional alternativo. Nesse caso se tem o aprendizado por meio de dados usados como exemplo para a solução de determinado problema. De acordo com Izbick e Santos (2020), as redes neurais artificiais podem ser entendidas como um estimador não-linear formado por camadas de entrada processamento e saída, conforme representado pela Figura 20.

Figura 20 — Representação de RNA com uma camada oculta.

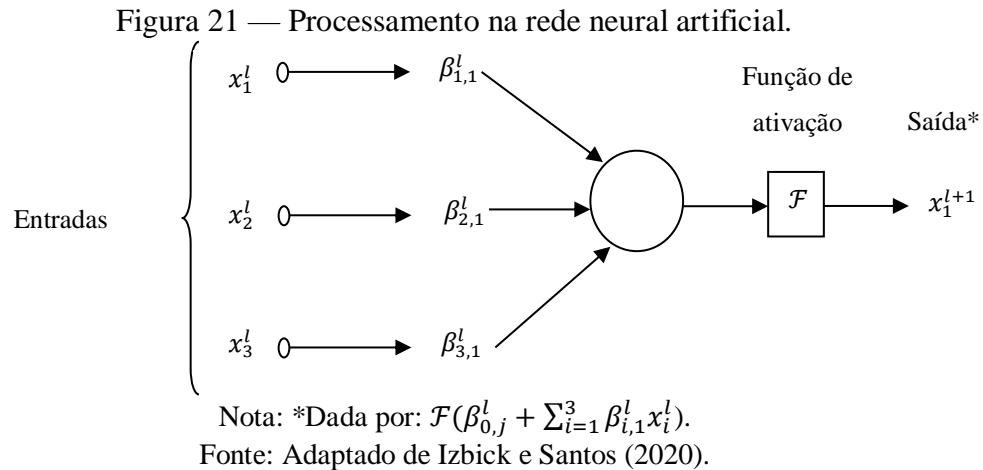


Fonte: Izbick e Santos (2020, p. 92).

Conforme se verifica na Figura 20, a rede neural é formada pela camada de entrada, camada oculta (que pode ser uma ou mais camadas) e camada de saída. A camada de entrada é formada pelas entradas da rede, que são as variáveis que compõem o banco de dados usado na análise. A camada oculta é onde o processamento é realizado, sendo que existem nós nessa camada, que transformam as informações (nós) da camada anterior.

Ademais as flechas representam pesos, que são adaptados para a execução da classificação (BRAGA; LUDERMIR; CARVALHO, 2000; IZBICK; SANTOS, 2020). Por

fim, está a camada de saída, que apresenta o resultado da classificação. A Figura 21 mostra de maneira mais detalhada o processamento em uma rede neural.



A técnica de redes neurais artificiais conta também com uma função de ativação. Existem diversos tipos de funções de ativação, como identidade, logística, tangente hiperbólica, *ReLU*, *Leaky ReLu* (IZBICK; SANTOS, 2020) e *sigmoide*, sendo que esta última é uma das mais usadas (AMBRÓSIO, 2002). A função de ativação, de acordo com Braga, Ludermir e Carvalho (2000), é responsável por ativar ou não a saída dos nós. Essa ativação dependerá do somatório do valor ponderado (pelos pesos) das entradas.

Sendo assim, a ativação dependerá da soma do produto das entradas e pesos da camada oculta. Entre os trabalhos que usaram a técnica de redes neurais artificiais, podem-se citar os de Obermann e Waack (2015), Barboza, Kimura e Altman (2017), Prado *et al.* (2020), Pinto (2021) e Santos (2021).

2.2.5 Máquina de vetores de suporte (*Support vector machines -SVMs*)

A técnica de Máquina de vetores de suporte (*Support vector machines - SVM*) foi desenvolvida por Cortes e Vapnik (1995). De acordo com os autores, inicialmente foi desenvolvida a técnica *support-vector networks* que era específica para a separação de dados sem erros pelos vetores de treinamento. Os autores ampliaram essa técnica, para que ela abrangesse casos em que houvesse a impossibilidade da separação sem erros nos vetores de treinamento. Ou seja, os autores ampliaram a técnica para casos em que os dados de treinamento que possuem erro.

A técnica de aprendizado de máquina foi uma nova opção para problemas envolvendo a classificação de dois grupos. A principal ideia foi o mapeamento não linear dos vetores de

entrada para um espaço com alta dimensão, no qual uma superfície de decisão linear seria construída.

Segundo Izbick e Santos (2020), a técnica de SVM se diferencia de outras técnicas de aprendizado de máquina, uma vez que sua motivação é diferente das de outras. De acordo com os autores, no SVM probabilidades não são estimadas e o resultado indica classes estimadas para as novas observações. A representação do SVM, de acordo com Izbick e Santos (2020) pode ser dada considerando uma função linear, representada pela Equação 5.

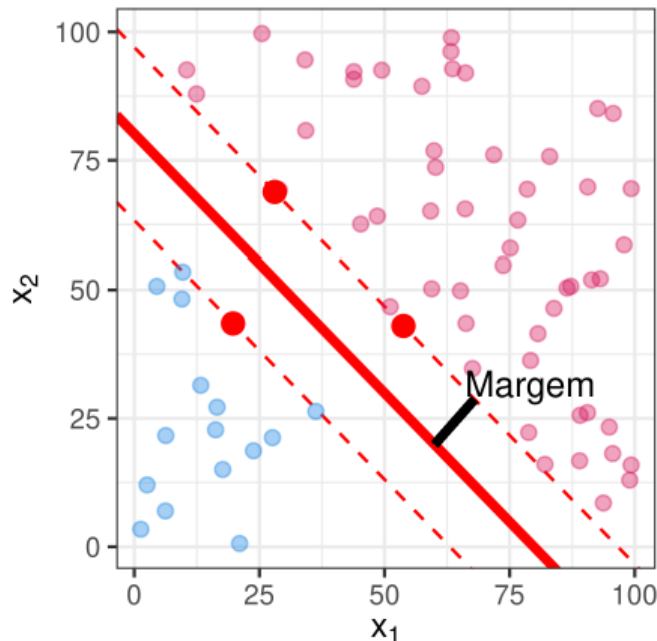
$$f(\mathbf{x}) = \beta_0 + \beta_1 x_1 + \dots + \beta_d x_d \quad (5)$$

Sendo que o classificador é obtido considerando que, se $f(\mathbf{x}) < 0$, $g(\mathbf{x}) = -1$; e se $f(\mathbf{x}) > 0$, $g(\mathbf{x}) = 1$.

De maneira complementar, para Mueller e Guido (2019) as SVMs são uma técnica que possibilita a construção de modelos mais complexos. Os autores afirmam que a SVM define o limite da decisão com base no aprendizado da importância de cada ponto que compõe os dados de treinamento.

O subconjunto que geralmente importa para tal decisão é o de pontos de dados que estão no limite entre as classes, que são chamados de vetores de suporte (responsáveis pela nomenclatura da técnica). Sendo assim, para a previsão de um novo ponto, é considerada a distância para cada vetor de suporte individualmente. A decisão se baseia nas distâncias até o vetor de suporte e na importância de cada vetor, que foi aprendida na etapa de treinamento. A Figura 22 mostra um exemplo da aplicação do SVM.

Figura 22 — Exemplo de aplicação do SVM.



Fonte: Izbick e Santos (2020, p. 153).

Conforme a Figura 22 mostrou, é possível perceber que a linha vermelha mostra a fronteira de decisão que permite a classificação dos dados, chamado hiperplano. O hiperplano, faz a separação das observações nos dados de treinamento (IZBICK; SANTOS, 2020). Os vetores de suporte estão em destaque na imagem, sendo os pontos com a cor vermelha e de tamanho maior que os demais. Os vetores de suporte são os pontos que estão mais próximos do hiperplano, e definem a margem, que é a distância entre os vetores de suporte e o hiperplano (IZBICK; SANTOS, 2020).

Entre os trabalhos que usaram o SVM estiveram os estudos de Obermann e Waack (2015), Barboza, Kimura e Altman (2017), Pinto (2021) e Santos (2021). Apesar de geralmente se assumir que técnicas de aprendizado de máquina de caixa preta são mais robustas, como o SVM, nos estudos mencionados outras técnicas apresentaram resultados mais robustos, entre elas técnicas de caixa branca (OBERMANN; WAACK, 2015) e técnicas de caixa cinza (BARBOZA; KIMURA; ALTMAN, 2017; SANTOS, 2021).

2.3 Curva ROC

Para a comparação entre os métodos de previsão a curva ROC foi usada. A curva ROC (*Receiver Operating Characteristic Curves*) se baseia nas dimensões de sensibilidade e especificidade (SICSÚ, 2010). A dimensão da sensibilidade é entendida como a dimensão capaz de gerar resultados positivos sobre uma determinada característica, enquanto a especificidade

pode ser entendida como a capacidade de geração de resultados negativos sobre uma determinada característica.

De acordo com Sicsú (2010), a dimensão da sensibilidade corresponde à capacidade de identificação de clientes ruins (no caso deste estudo, empresas ruins ou insolventes: verdadeiros positivos), enquanto a especificidade corresponde à capacidade de identificação de clientes bons (no caso deste estudo, empresas boas ou não insolventes: verdadeiros negativos).

A curva ROC apresenta a área sob a curva (AUC: *Area Under the Curve*), que é a principal medida que deve ser observada. A AUC varia de zero (0) a um (1), e quanto maior o valor da AUC melhor é o ajuste do modelo (SOARES; REBOUÇAS, 2014). Segundo Sumihara Filho e Slegers (2010), algumas classificações do ajuste do modelo são feitas com base no valor da AUC.

De acordo com Sumihara Filho e Slegers (2010), quando o valor da AUC é menor ou igual a 0,5, considera-se que o ajuste do modelo aos dados não existe; caso o valor da AUC esteja entre 0,5 e 0,7, considera-se que o ajuste do modelo aos dados é baixo; caso esteja entre 0,7 e 0,8 considera-se que o ajuste do modelo aos dados é aceitável; caso o valor da AUC esteja entre 0,8 e 0,9 o ajuste do modelo aos dados é considerado excelente. Por fim, caso o valor da AUC seja maior que 0,9, o ajuste do modelo aos dados é considerado acima do comum.

3 METODOLOGIA

Neste tópico a metodologia para a análise proposta no estudo está detalhada. Para isso, foram especificados os procedimentos de coleta de dados, como a amostra será definida, as variáveis que serão utilizadas e os procedimentos de análise. Cada uma dessas etapas está detalhada em seguida.

3.1 Procedimentos de coletados dados

Para a coleta dos dados, a base de dados Economatica foi utilizada. Por meio dela, foi possível ter acesso aos dados das demonstrações financeiras de empresas com ações negociadas na B3. Uma vez que para a análise será necessário o cálculo de indicadores econômico-financeiros, por meio dos dados disponíveis na base de dados Economatica esses indicadores puderam ser calculados.

O período de análise escolhido foi do ano de 2010 ao ano de 2021. O ano de 2010 foi escolhido como ano inicial devido à alteração da norma contábil brasileira, que objetivou a padronização da legislação brasileira (Lei das Sociedades anônimas – lei 6.404/76) às normas internacionais, por meio das leis 11.638/07 e 11.941/09 (FREIRE *et al.*, 2012). Dessa maneira, usando o período posterior à alteração da norma contábil será possível trabalhar com os dados já padronizados e segundo uma única norma.

3.2 Amostra

A amostra do estudo foi composta por empresas não-financeiras de capital aberto com ações negociadas na B3. As empresas financeiras foram excluídas para que a amostra se tornasse mais homogênea, pois as empresas financeiras trabalham com produtos e serviços financeiros, fato que faz com que elas apresentem características específicas, que as diferenciam das empresas não-financeiras.

Uma vez que o foco do estudo se concentrou na definição de modelos de previsão de insolvência, inicialmente foram selecionadas as empresas classificadas como insolventes, sendo elas empresas em recuperação judicial, recuperação extrajudicial, falidas ou ainda aquelas que apresentaram insuficiência de saldo (patrimônio líquido negativo) (PEDRO JÚNIOR, 2011; PRADO, 2016; SILVA; FONSECA; LITTIG, 2019; STÜPP, 2015).

Adicionalmente, a literatura sugere que estudos de previsão apresentem também empresas solventes na amostra (amostra equiparada), tendo ao menos uma empresa solvente para cada empresa insolvente, de preferência do mesmo setor e com um valor do ativo total próximo (BRITO; ASSAF NETO; CORRAR, 2009; PRADO, 2016; SANVICENTE; MINARDI, 1998). Destaca-se também que os dados que compuseram a amostra foram dados anuais, referentes a um ano antes do ano em que a empresa entrou em situação de insolvência, uma vez que pretende-se prever quais as empresas insolventes com certa antecedência (PRADO, 2016).

Para a seleção das empresas em recuperação judicial, extrajudicial ou falência, serão usadas empresas que compuseram as amostras dos estudos de Stüpp (2015), Prado (2016) e Silva, Fonseca e Littig (2019), que foram trabalhos anteriores de previsão de insolvência que usaram empresas brasileiras de capital aberto. Ademais, a amostra será atualizada, ou seja, serão adicionadas empresas em situação de recuperação judicial ou extrajudicial conforme divulgação atual do site da B3.

Sendo assim, para a definição da amostra de empresas insolventes, as empresas das amostras dos estudos anteriores foram listadas, bem como as empresas que se encontram atualmente em recuperação judicial ou extrajudicial, de acordo com o site da B3. Após a listagem, foram excluídas da amostra empresas repetidas em mais de uma listagem.

Para a seleção de empresas com insuficiência de saldo, foram filtradas as empresas que apresentaram patrimônio líquido negativo pelo menos uma vez durante o período analisado (2010 a 2021). Após essa filtragem, a lista de empresas com patrimônio líquido negativo e empresas em situação de recuperação judicial, extrajudicial ou falência foram comparadas, a fim, novamente de excluir empresas que estivessem repetidas. Assim, foi possível definir a lista final de empresas insolventes sem duplicidade.

Conforme mencionado, foram necessárias empresas solventes na amostra do estudo. Uma vez que os estudos de Stüpp (2015), Prado (2016) e Silva, Fonseca e Littig (2019), apresentaram em suas amostras uma empresa solvente para cada empresa insolvente, dando preferência a empresas do mesmo setor e com o valor do ativo total mais próximo, neste estudo foram considerados estes mesmos parâmetros para a seleção da amostra equiparada.

Para a amostra equiparada, foram selecionadas empresas preferencialmente do mesmo setor e com valor mais próximo do ativo total da empresa insolvente correspondente. Para essa seleção foram selecionadas empresas do mesmo setor de atuação, considerando o setor Econômica, e o valor mais próximo do ativo total ajustado por inflação do mês de dezembro do ano em que a empresa insolvente correspondente entrou em situação de insolvência

(recuperação judicial, extrajudicial ou falência). Na Tabela 2 encontram-se as empresas em recuperação judicial durante o período analisado (2010 a 2021).

Tabela 2 — Empresas em recuperação judicial e empresas correspondentes (continua).

Empresa em Recuperação judicial	Ano do Evento	Ativo dezembro (Reais Mil)	Setor (econômica)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Bardella	2019	877.775	Máquinas Indust	Aeris	966.860
Brasil Pharma S/A	2018	90.712	Comércio	Enjoei	65.490
Buettner S/A	2011	391.219	Textil	Encorpar PN	410.846
Clarion S/A Agroindustrial	2013	1.971.441	Alimentos e Beb	Josapar	2.256.054
Dexxos Participações S/A	2013	1.169.759	Outros	Time For Fun	1.269.677
Dhb	2015	377.174	Veiculos e peças	Recrusul	50.256
Eneva S/A	2014	11.127.233	Energia Elétrica	TranPaulist	11.281.591
Equatorial On	2012	16.792.918	Energia Elétrica	Light Serv de Eletr S/A	15.964.963
Eternit	2018	715.559	Minerais não Met	Portobello	2.039.491
Fabrica Carlos Renaux	2011	275.503	Textil	Arezzo	808.874
Fer Heringer	2019	1.904.402	Química	Ourofino S/A	1.048.112
Fibam Companhia Industrial	2014	136.755	Siderur & Metalur	Sid Nacional	78.611.198
Hotéis Othon	2018	606.620	Outros	Ambipar	574.940
Inepar	2014	4.218.279	Outros	Aegea Saneamento e Part S/A	4.241.208
Joao Fortes	2020	1.476.618	Construção	Mitre Realty	1.523.692
LarkMaqs	2012	18.253	Outros	Minasmaquinas	198.982
Lupatech	2014	1.615.326	Siderur & Metalur	Ferbasa	2.308.604
Mangels Indl	2013	754.029	Siderur & Metalur	Panatlantica	978.906
Metalúrgica Duque S/A	2014	248.743	Siderur & Metalur	Aliperti	577.271
Mmx Miner	2014	5.114.451	Mineração	Litel	50.089.859
Oi S/A	2016	110.340.158	Telecomunicações	Telef Brasil	137.055.410
Osx Brasil	2013	14.358.364	Veiculos e peças	Iochnp-Maxion	9.934.026
Pdg Realt S/A	2017	3.872.484	Construção	Gafisa	3.176.223
Pet Manguinh	2013	347.618	Petróleo e Gas	Petroleo Lub do Nordeste S/A	29.235

Tabela 2 — Empresas em recuperação judicial e empresas correspondentes (conclusão).

Empresa em Recuperação judicial	Ano do Evento	Ativo dezembro (Reais Mil)	Setor (econômica)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Pomifrutas S/A	2018	36.137	Agro e Pesca	Agribrasil	29.770
Rede Energia	2012	5.025.600	Energia Elétrica	Coelce	6.338.181
Renova	2019	2.760.346	Energia Elétrica	Statkraft	2.656.049
Saraiva Livr	2018	1.230.460	Outros	Mills	1.275.689
Schlosser	2015	160.552	Textil	IndCataguas	384.628
Sultepa	2015	1.688.603	Construção	Rni	2.434.271
Teka	2012	1.490.833	Textil	Cedro	1.206.083
Viver	2016	1.321.859	Construção	Trisul	1.169.673
Wetzel S/A	2016	262.719	Veiculos e peças	Bic Monark	290.738

Fonte: Da autora (2023).

Na Tabela 3, por sua vez, estão as empresas em recuperação extrajudicial do estudo, juntamente com as respectivas empresas solventes.

Tabela 3 — Empresas em recuperação extrajudicial e empresas correspondentes.

Empresa em Recuperação Extrajudicial	Ano do Evento	Setor (econômica)	Ativo dezembro (Reais Mil)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Atmasa	2020	Outros	1.849.516	Saneatins - Cia De Saneamento Do Tocantins	1.822.160
Le Lis Blanc	2020	Textil	1.594.263	Springs	3.349.122
Nutriplant	2017	Química	90.864	Metanor S/A Metanol Do Ne	139.827
Tex Renaux	2019	Textil	222.215	Dohler	894.517
Triunfo	2017	5.985.773,5072	Transport e Serviç	Hidrovias	4.808.151,9180

Fonte: Da autora (2023).

Por fim, estiveram as empresas em situação de insolvência de saldo, ou seja, patrimônio líquido negativo. Na Tabela 4 estão as empresas com patrimônio líquido negativo e as empresas solventes correspondentes.

Tabela 4 — Empresas com patrimônio líquido negativo e correspondentes (continua).

PL Negativo	Data (Ano)	Ativo dezembro (Reais Mil)	Setor (econômica)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Agpart	2017	2.804.331	Outros	Ser Educa	2.750.311
All Am Lat Log Oeste S/A	2013	1.413.015	Transporte Serviç	Autopista Fluminense S/A	1.383.511

Tabela 4 — Empresas com patrimônio líquido negativo e correspondentes (continua).

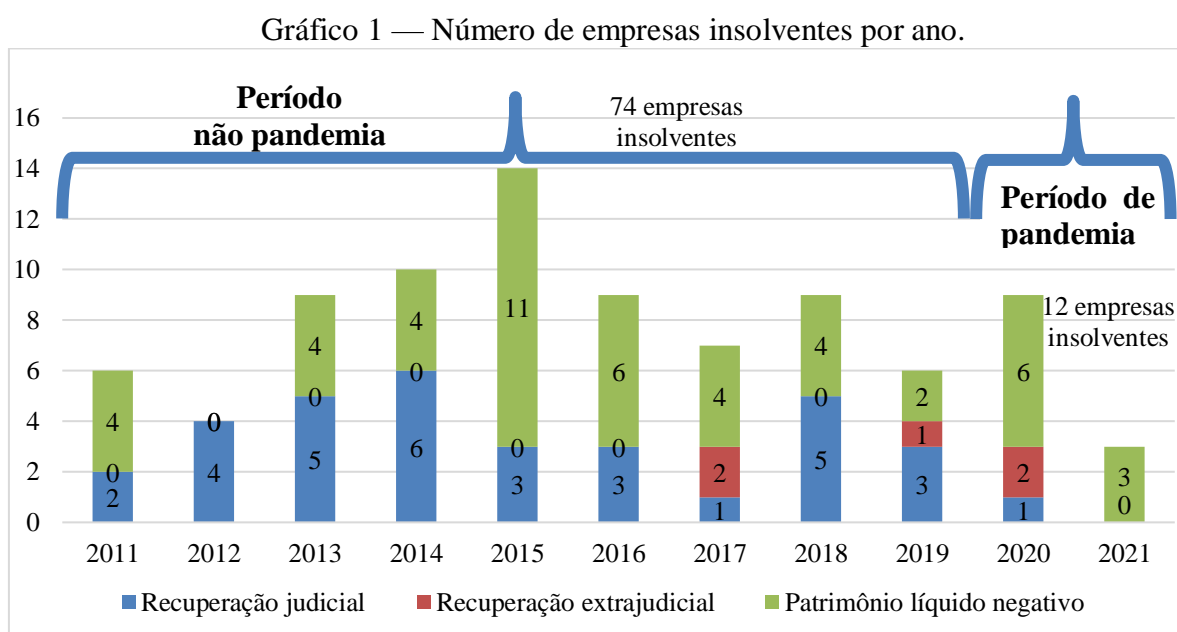
PL Negativo	Data (Ano)	Ativo dezembro (Reais Mil)	Setor (econômica)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Alphaville	2018	3.241.779	Construção	Tenda	3.295.361
Altus S/A	2016	145.794	Siderur & Metalur	Tekno	294.883
Azevedo	2017	174.024	Construção	Priner	237.923
Azul	2019	23.140.067	Transporte Serviço	Rumo S.A.	39.575.488
B Tech Eqi	2016	122.472	Alimentos e Beb	Excelsior	93.954
Biosev	2015	14.638.798	Outros	Sanepar	11.767.168
Br Home	2016	349.868	Construção	Cr2	259.947
Brasília Participações S/A	2020	17.580	Energia Elétrica	Afluentes T	239.630
Braskem	2020	99.278.187	Química	Unipar	5.199.293
Cambuci	2015	381.873	Textil	Pettenati	651.281
Carbomil S/A	2011	91.850	Mineração	Vale	493.183.608
Cee-D	2015	4.732.692	Energia Elétrica	Ceb	4.773.745
Cinesystem	2020	213.103	Outros	Trevisa	176.809
Cobrasma	2015	232.415	Veículos e peças	Fras-Le	1.385.796
Conc do Aerop Inter de Guarulhos S/A	2015	23.584.335	Transporte Serviço	CCR AS	30.947.652
Concess Rodovias Tietes	2018	2.026.685	Transporte Serviço	Ecopistas - Concess Das Rodov Ayrtton Senna S/A	2.028.729
Concessionaria BR-040 S/A	2020	1.201.986	Transporte Serviço	Viarondon Conc de Rodov S/A	1.222.841
Conpel Cia Nordestina Papel	2013	155.791	Outros	Marina de Iracema Park S/A	166.259
Ferrovias Sul Atlântico S/A	2016	4.236.860	Transporte Serviço	Concess Rota Das Bandeiras S/A	4.474.826
Flex S/A	2021	424.068	Outros	Self It Academias Holding S/A	442.171
Gazola	2011	27.303	Siderur & Metalur	Metisa	468.160
Gol	2014	15.758.928	Transporte Serviço	JSL	12.477.822
Haga	2014	74.962	Siderur & Metalur	Aco Altona	431.228
Iguacu Café	2011	1.109.113	Alimentos e Beb	J. Macedo	1.195.794

Tabela 4 — Empresas com patrimônio líquido negativo e correspondentes (conclusão).

PL Negativo	Data (Ano)	Ativo dezembro (Reais Mil)	Setor (econômica)	Empresa Solvente (Correspondente)	Ativo dezembro (Reais Mil)
Invepar	2019	28.713.155	Transporte Serviço	Ecorodovias	13.912.919
Karsten	2013	611.864	Textil	Santanense	700.733
Log-In	2017	1.528.701	Transporte Serviço	Vix Logistica S/A	1.532.245
Mendes Jr	2016	1.288.523	Construção	You Inc Incorp e Part S/A	974.736
Metal Iguaçu	2020	88.493	Siderur & Metalur	Gerdau	72.798.009
Metalrio	2015	1.720.367	Máquinas Indust	Romi	1.740.833
Minerva	2015	11.863.326	Alimentos e Beb	Marfrig	29.852.177
Minupar	2014	320.334	Alimentos e Beb	Oderich	604.516
Mundial	2015	1.266.870	Siderur & Metalur	Kepler Weber	1.129.285
Padtec	2015	488.120	Software e Dados	Sinqia	140.161
Paranapanema	2020	4.649.663	Siderur & Metalur	Gerdau Met	74.691.066
Plascar Part	2016	780.732	Plascar	Metal Leve	3.162.197
Portx	2011	2.889.828	Transporte Serviço	Tegma	1.795.772
PqHopiHari	2013	512.529	Outros	Seiva S/A - Florestas e Indústrias	497.853
Rossi Resid	2018	3.196.297	Construção	Tecnisa	2.434.709
Rumo Malha Paulista S/A	2017	7.216.440	Transporte Serviço	Autoban - Conc do Sist Anhanguera Bandeirantes S/A	5.310.912
Salus Infra Protuaria S/A	2018	457.448	Outros	Espacolaser	974.956
Snb Participacoes S/A	2021	8.695	Outros	Hauscenter S/A	4.970
SPturis	2021	328.578	Outros	Hospital Anchieta S/A	325.460
Tectoy	2014	113.361	Outros	Santa Catarina Part Invest S/A	110.965
Telebras	2015	3.645.938	Telecomunicações	Algar Telecom S/A	4.537.649
Termeletrica Pernambuco 3 AS	2015	809.931	Energia Elétrica	Omega Geração S/A	1.096.018

Fonte: Da autora (2023).

Diante disso, percebe-se que a amostra proposta para este estudo foi composta por 172 empresas não financeiras com ações negociadas na B3. Destas, 86 são empresas solventes e 86 são empresas insolventes. Dentro da amostra de empresas insolventes, 33 empresas estiveram em recuperação judicial, cinco em recuperação extrajudicial e 48 apresentaram patrimônio líquido negativo. Das 86 empresas insolventes, 74 entraram em situação de insolvência antes do período de pandemia de COVID-19, e 12 delas no período de pandemia, ou seja, nos anos de 2020 e 2021, conforme mostrado no Gráfico 1.



Fonte: Da autora (2023).

Uma vez que para estudos de previsão faz-se necessário uma amostra de treinamento e pretendeu-se analisar se os resultados durante a pandemia de COVID-19 são satisfatórios, neste estudo utilizou-se como amostra de treinamento as empresas que se tornaram insolventes no período de 2010 a 2019 e suas empresas solventes correspondentes.

Para a amostra de teste, portanto, foram usadas as empresas que se tornaram insolventes durante o período de pandemia, ou seja, nos anos de 2020 e 2021, e suas empresas solventes correspondentes. Sendo assim, a amostra de treinamento correspondeu a um percentual de aproximadamente 86,5% (148 empresas) e a de validação de aproximadamente 13,5% (24 empresas). Para isso, o banco de dados foi organizado em ordem alfabética e, posteriormente em ordem cronológica, para que as empresas que apresentaram situação de insolvência nos anos de 2020 e 2021 estivessem por último no banco de dados, permitindo que ao definir as amostras

de treinamento e teste, as empresas insolventes no período de pandemia fossem definidas como a amostra de teste.

3.3 Variáveis - Indicadores Econômico-financeiros

A variável dependente do estudo foi uma variável dicotômica, do tipo nominal e não-métrica (HAIR *et al.*, 2006). A variável se classifica dessa forma pois indica se a empresa se classifica como insolvente (sim ou não). Ou seja, irá apresentar o valor de um (1) caso a empresa seja insolvente e zero (0) caso a empresa não seja insolvente (solvente).

As variáveis utilizadas na análise foram obtidas por meio da seleção de indicadores econômico-financeiros presentes em estudos de previsão anteriores. Entre os estudos estão os de Altman (1968), Kanitz (1974), Elizabetsky (1976), Matias (1978), Altman, Baydia e Dias (1979), Silva (1982), Sanvicenti e Minardi (1998), Akiama (2008), Stüpp (2015) e Prado *et al.* (2020).

Inicialmente, 44 variáveis foram listadas, as quais foram numeradas de X_1 a X_{44} . No entanto, devido às alterações nas normas contábeis brasileiras para a padronização de acordo com as normas internacionais, algumas contas foram excluídas, fato que impossibilitou o cálculo de alguns desses indicadores. Ademais, em alguns casos a ocorrência de dados faltantes também fez com que alguns dos indicadores econômico-financeiros fossem desconsiderados. Diante disso, após a exclusão dos indicadores necessários, o total de variáveis iniciais do estudo foi de 27.

Para a seleção de um número reduzido de variáveis para os modelos de previsão, utilizou-se a técnica de floresta aleatória, por meio da qual é possível determinar a ordem de importância das variáveis, o que possibilita a seleção de um número reduzido de variáveis, priorizando aquelas com maior nível de importância. O Quadro 3 detalha as variáveis usadas nos estudos de previsão anteriores e que foram usados no presente estudo.

Quadro 3 — Indicadores econômico-financeiros de estudos anteriores (continua).

Nome	Variável	Trabalhos em que apareceram
X_1	$\frac{\textit{Ativo Circulante}}{\textit{Passivo Circulante}}$	Kanitz (1974), Matias (1978), Silva (1982), Stüpp (2015).
X_2	$\frac{\textit{Receita líquida operacional}}{\textit{Ativo Total}}$	Altman (1968), Altman, Baidya e Dias (1979), Stüpp (2015).

Quadro 3 — Indicadores econômico-financeiros de estudos anteriores (continua).

Nome	Variável	Trabalhos em que apareceram
X_3	$\frac{\text{Caixa e equivalente de caixa}}{\text{Ativo Total}}$	Matias (1978), Silva (1982), Prado et al. (2020).
X_3	$\frac{\text{Caixa e equivalente de caixa}}{\text{Ativo Total}}$	Matias (1978), Silva (1982), Prado et al. (2020).
X_5	$\frac{\text{Lucro Antes de Juros e Impostos}}{\text{Ativo Total}}$	Altman (1968), Altman, Baidya e Dias (1979).
X_6	$\frac{\text{Passivo Circulante} + \text{Passivo não Circulante}}{\text{Ativo Total}}$	Silva (1982), Stüpp (2015).
X_7	$\frac{\text{Lucro Líquido}}{\text{Patrimônio Líquido}}$	Kanitz (1974), Stüpp (2015).
X_8	$\frac{\text{Passivo Circulante} + \text{Passivo Não Circulante}}{\text{Patrimônio Líquido}}$	Kanitz (1974), Stüpp (2015).
X_9	$\frac{\text{Ativo Circulante} + \text{Realizável a Longo Prazo}}{\text{Passivo Circulante} + \text{Passivo não Circulante}}$	Kanitz (1974), Stüpp (2015).
X_{10}	$\frac{\text{Caixa e equivalente de caixa}}{\text{Ativo não circulante} - \text{Realizável a Longo Prazo}}$	Elisabetsky (1976), Silva (1982).
X_{11}	$\frac{\text{Estoque}}{\text{Ativo Total}}$	Elisabetsky (1976), Silva (1982).
X_{12}	$\frac{\text{Fornecedores}}{\text{Ativo Total}}$	Matias (1978), Silva (1982).
X_{13}	$\frac{\text{Patrimônio Líquido}}{\text{Ativo Total}}$	Matias (1978), Akiama(2008).
X_{14}	$\frac{\text{Ativo Circulante} - \text{Estoques}}{\text{Passivo Circulante}}$	Kanitz (1974).
X_{15}	$\frac{\text{Patrimônio Líquido}}{\text{Passivo Total}}$	Altman, Baidya e Dias (1979), Sanvicente e Minardi (1998).
X_{16}	$\frac{\text{Ativo Não Circulante}}{\text{Patrimônio Líquido}}$	Stüpp (2015), Akiama (2008).

Quadro 3 — Indicadores econômico-financeiros de estudos anteriores (conclusão).

Nome	Variável	Trabalhos em que apareceram
X_{17}	$\frac{\textit{Ativo Circulante} - \textit{Passivo Total}}{\textit{Ativo Total}}$	Sanvicente e Minardi (1998).
X_{21}	$\frac{\textit{Capital de Giro}}{\textit{Ativo Total}}$	Altman (1968).
X_{22}	$\frac{\textit{Capital Circulante Líquido}}{\textit{Ativo Total}}$	Akiama (2008).
X_{23}	$\frac{\textit{Financiamentos e Empréstimos Bancários}}{\textit{Ativo Circulante}}$	Matias (1978).
X_{24}	$\frac{\textit{Lucro Antes de Juros e Impostos} - \textit{Despesas Financeiras} + \textit{Receitas Financeiras}}{\textit{Ativo Total}}$	Sanvicente e Minardi (1998).
X_{27}	$\frac{\textit{Passivo Circulante}}{\textit{Ativo Total}}$	Elisabetsky (1976).
X_{28}	$\frac{\textit{Patrimônio Líquido} - \textit{Capital Social}}{\textit{Ativo Total}}$	Sanvicente e Minardi (1998)
X_{31}	$\frac{(\textit{Ativo Total} - \textit{Obrigações sociais e trabalhistas, Impostos a pagar})}{\textit{Patrimônio Líquido}}$	Silva (1982)
X_{32}	$\frac{\textit{Ativo não circulante} - \textit{Realizável a Longo Prazo}}{\textit{Patrimônio Líquido}}$	Akiama (2008)
X_{39}	$\frac{\textit{Lucro Antes de Juros e Impostos}}{\textit{Despesas Financeiras}}$	Sanvicente e Minardi (1998)
Nome	Variável	Trabalhos em que apareceram
X_{40}	$\frac{\textit{Lucro Antes dos Juros e Impostos}}{\textit{Dívida Líquida}}$	Stüpp (2015)
X_{44}	$\frac{\textit{Passivo Circulante}}{\textit{Passivo Circulante} + \textit{Passivo Não Circulante}}$	Stüpp (2015).

Fonte: Da autora, com base em Altman (1968), Kanitz (1974), Elisabetsky (1976, *apud* REBELLO, 2010), Matias (1978), Altman, Baydia e Dias (1979), Silva (1982, *apud* MÁRIO, 2002), Sanvicenti e Minardi (1998), Akiama (2008), Stüpp (2015) e Prado *et al.* (2020).

Após a definição das variáveis, realizou-se a etapa de winsorização dos dados, com o intuito de remover *outliers* que poderiam comprometer a amostra. Para isso a winsorização foi realizada considerando o valor máximo de 5% dos dados (HAIR, *et al.*, 2006), procurando e chegar a valores de assimetria próximos de zero e valores de curtose próximos de três (FÁVERO; BELFIORE, 2017). Sendo assim, novas variáveis foram criadas, e as variáveis winsorizadas receberam em sua nomenclatura o sufixo “w” (exemplo: variável X_1 winsorizada chamou-se de X_{1_w}).

3.4 Procedimentos de análise

Para análise do estudo, diferentes técnicas de previsão de insolvência foram utilizadas e comparadas. Conforme já mencionado, diversos estudos de previsão já foram realizados, considerando diferentes técnicas de previsão. De maneira geral, as técnicas estatísticas foram muito usadas principalmente nos anos iniciais de desenvolvimento de estudos de previsão e, posteriormente, técnicas de inteligência artificial, em especial as de aprendizado de máquina passaram a ser usadas (PRADO *et al.*, 2016).

No estudo, utilizaram-se duas técnicas estatísticas, que foram a análise discriminante e a regressão logística, as técnicas estatísticas mais utilizadas até então (PRADO *et al.*, 2016). De maneira complementar, técnicas de aprendizado de máquina de diferentes classificações também foram utilizadas. Tal escolha possibilitou verificar se as técnicas de caixa branca apresentam resultados similares às técnicas de caixa preta, uma vez que as técnicas de caixa branca geralmente são consideradas como mais interpretáveis e menos robustas, mas tem apresentado resultados satisfatórios e em alguns casos superiores às técnicas de caixa preta (OBERMANN; WAACK, 2015).

As técnicas de aprendizado de máquina foram árvore de decisão (classificada como técnica de caixa branca), K-vizinhos mais próximos (classificada como técnica de caixa branca), floresta aleatória (classificada como técnica de caixa cinza), RNAs e SVMs (classificadas como técnicas de caixa preta).

Como o estudo pretendeu realizar a comparação dos métodos, observou-se a acurácia e a curva ROC (*Receiver Operating Characteristic Curves*) de cada um dos modelos. A acurácia foi obtida pela divisão entre a quantidade de acertos da técnica em relação ao total de valores preditos. Com relação à curva ROC, observou-se a área sob a curva (AUC: *Area Under the Curve*), medida indica o nível de ajuste do modelo (quanto maior melhor) (SOARES; REBOUÇAS, 2014).

Ademais, observou-se também outras medidas de desempenho que são usadas na literatura (BARBOZA; KIMURA; ALTMAN, 2017; KIM; UPNEJA, 2014; WANG *et al.*, 2012), sendo elas a taxa de verdadeiro positivo (TVP) (1 - erro tipo I) e a taxa de verdadeiro negativo (TVN) (1 - erro tipo II). De acordo com Barboza, Kimura e Altman (2017), o erro tipo I pode ser obtido subtraindo de 1 a divisão do total de verdadeiros positivos, pela soma dos totais de verdadeiros positivos e falsos negativos, enquanto o erro tipo II pode ser obtido subtraindo de 1 a divisão do total verdadeiros negativos pela soma dos totais de verdadeiros negativos e falsos positivos.

Para a operacionalização das técnicas estatísticas e de aprendizado de máquinas utilizou o *Google Colab*. Por meio dele, o uso das ferramentas do *Python* é possível, sem a necessidade de instalação do *software*.

4 RESULTADOS

Por meio da técnica de floresta aleatória (RF), foi possível definir a ordem de importância dos indicadores econômico-financeiros, considerando as 27 inicialmente usadas. A partir disso, os 10 indicadores econômico-financeiros mais relevantes foram X_{6-w} , X_{17-w} , X_{24-w} , X_{28-w} , X_{15-w} , X_{39-w} , X_{13-w} , X_{40-w} , X_{14-w} e X_{1-w} .

No entanto, de acordo com Barboza, Kimura e Altman (2017), é necessário analisar se existe correlação alta entre as variáveis, principalmente devido ao uso das técnicas estatísticas de análise discriminante e regressão logística. Para isso, analisaram-se os valores do fator de inflação da variância (VIF). De acordo com Gujarati e Porter (2011), valores do VIF superiores a 10 indicam a existência de alta multicolinearidade, assim, as variáveis do modelo que apresentaram valores do VIF superiores a 10 foram excluídas do modelo.

As variáveis que estiveram entre as 10 principais, mas que foram excluídas devido à alta correlação foram X_{6-w} , X_{17-w} , X_{13-w} e X_{1-w} . Após a exclusão das variáveis que indicavam a existência de multicolinearidade, o total de variáveis usadas nos modelos de previsão foi de seis, que estão representadas na Tabela 5, que mostrou também os valores do VIF de cada uma delas.

Tabela 5 — Indicadores econômico-financeiros usados nos modelos preditivos e VIF.

Nome	Variável	VIF
X_{14}	$\frac{\text{AtivoCirculante} - \text{Estoques}}{\text{PassivoCirculante}}$	1.67
X_{15}	$\frac{\text{PatrimônioLíquido}}{\text{Passivo Total}}$	2.11
X_{24}	$\frac{\text{LucroAntesdeJuroseImpostos} - \text{DespesasFinanceiras} + \text{ReceitasFinanceiras}}{\text{AtivoTotal}}$	2.79
X_{28}	$\frac{\text{PatrimônioLíquido} - \text{Capital Social}}{\text{AtivoTotal}}$	2.58
X_{39}	$\frac{\text{LucroAntesdeJuroseImpostos}}{\text{DespesasFinanceiras}}$	3.83
X_{40}	$\frac{\text{LucroAntesdosJuroseImpostos}}{\text{DívidaLíquida}}$	5.07

Fonte: Da autora (2023).

As estatísticas descritivas das variáveis usadas nos modelos de previsão foram detalhadas na Tabela 6, sendo elas média, mínimo, máximo, desvio padrão, variância, assimetria e curtose.

Tabela 6 — Estatísticas descritivas dos indicadores econômico-financeiros.

Var.	Média	Mínimo	Máximo	D. Padrão	Variância	Assimetria	Curtose
X_{14-w}	0.001	1.26e-07	0.004	0.0009	9.90e-07	1.8134	6.226
X_{15-w}	0.707	-0.993	6.116	1.458	2.127	2.695	10.329
X_{24-w}	-0.079	-1.168	0.176	0.252	0.0637	-2.865	12.483
X_{28-w}	116.543	-1785.87	994.808	601.299	361559.3	-1.971	6.875
X_{39-w}	0.626	-9.480	13.015	3.832	14.684	0.733	6.142
X_{40-w}	0.000	-0.0006	0.0007	0.0002	4.85e-08	-0.032	5.004

Fonte: Da autora (2023).

A partir dessas variáveis, as técnicas de previsão foram executadas. Conforme já mencionado, as técnicas comparadas neste estudo foram técnicas estatísticas (análise discriminante e regressão logística) e técnicas de aprendizado de máquinas de caixa branca (K-vizinhos mais próximos e árvore de decisão), de caixa cinza (floresta aleatória) e de caixa preta (redes neurais e máquinas de vetores de suporte). A partir dos resultados, foram analisadas as medidas de acurácia, verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP), falsos negativos (FN), erro tipo I, erro tipo II e a AUC da amostra de teste de cada uma das técnicas analisadas, que foram detalhadas na Tabela 7:

Tabela 7 — Resultados das técnicas de previsão.

Modelo	Acurácia	VP	VN	FP	FN	Erro tipo 1	Erro tipo 2	AUC
LDA	70.83%	7	10	2	5	41.67%	16.67	0.7083
LR	70.83%	8	9	3	4	33.33%	25%	0.7569
KNN	66.67%	7	9	3	5	41.67%	25%	0.7917
DT	79.17%	8	11	1	4	33.33%	8.33%	0.7917
RF	75%	8	10	2	4	33.33%	16.67%	0.75
RNA	66.67%	7	9	3	5	41.67%	25%	0.7986
SVM	66.67%	7	9	3	5	41.67%	25%	0.7083

Fonte: Da autora (2023).

Ao analisar a acurácia dos modelos de previsão, a técnica que apresentou o melhor percentual de acertos foi a de árvore de decisão (DT), com 79,17%, em seguida esteve a de floresta aleatória (RF), com 75%, regressão logística e análise discriminante, com 70,83% e,

por fim, estiveram a rede neural artificial (RNA), k-vizinhos mais próximos (KNN) e máquina de vetores de suporte (SVM), com 66,67%.

Esse resultado mostrou que, considerando as empresas que entraram em situação de insolvência como a amostra de teste, as técnicas de caixa branca e caixa cinza apresentaram as maiores acurácias, enquanto as técnicas de caixa preta (RNA e SVM) apresentaram os valores de acurácia mais baixos.

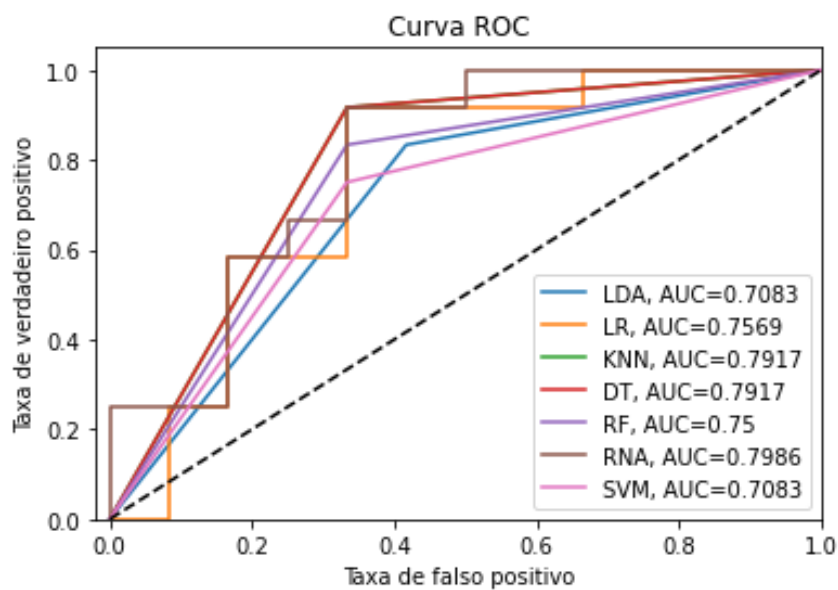
Esses resultados corroboram os achados de Obermann e Waack (2015), que mostraram a não inferioridade de modelos de caixa branca em relação aos modelos de caixa cinza e caixa preta. Os resultados também corroboram parcialmente os achados de Barboza, Kimura e Altman (2017), em que as técnicas de caixa preta (RNA e SVM) não superaram técnicas mais interpretáveis, como a de RF, que segunda técnica com melhor acurácia neste estudo (acurácia = 75%).

Ademais, menciona-se a não superioridade dos algoritmos de caixa preta em relação às técnicas estatísticas convencionais. As técnicas convencionais de análise discriminante e regressão logística, apresentaram acurácia também superiores às das técnicas de RNA e SVM. Esses resultados contradizem os estudos de Wilson e Sharda (1994), West (2000) e Baesens *et al.* (2003), nos quais as técnicas de aprendizado de máquina (RNA e/ou SVM) superaram os resultados das técnicas estatísticas tradicionais (LDA e/ou LR).

Tais resultados mostraram que, considerando uma amostra composta por empresas brasileiras, a capacidade de previsão diante de um período de crise (especificamente do período de crise causado pela pandemia de COVID-19), as técnicas de aprendizado de máquinas mais interpretáveis e técnicas estatísticas tradicionais apresentaram resultados superiores às técnicas menos interpretáveis. Esse resultado contradiz o argumento de que as técnicas de caixa preta são geralmente mais robustas e menos interpretáveis, se comparadas às técnicas de caixa branca e caixa cinza.

Na Figura 23 encontra-se a representação gráfica da curva ROC dos modelos analisados no estudo. Conforme mencionado anteriormente, a curva ROC mostra o ajuste dos modelos aos dados.

Figura 23 — Curva ROC de todos os modelos de previsão.



Fonte: Da autora (2023).

Por meio da Figura 23, percebe-se que as técnicas com maior AUC foram a de RNA (AUC=0,7986), KNN e DT (ambas com AUC=0,7917) e LR (AUC=0,7569). Em seguida estiveram as técnicas de RF (AUC=0,75) e LDA e SVM (ambas com AUC=0,7083). De acordo com Sumihara Filho e Slegers (2010), quando a AUC está entre 0,7 e 0,8, o ajuste do modelo aos dados é aceitável. Assim, como todas as técnicas analisadas apresentaram valores maiores que 0,7 e menores que 0,8, considera-se que o ajuste de todas as técnicas pode ser considerado como aceitável.

5 CONSIDERAÇÕES FINAIS

Os estudos envolvendo a previsão de insolvência são considerados relevantes face às possíveis consequências da falência de uma empresa, que pode causar um efeito sistêmico (BATTISTON *et al.*, 2007). Esse efeito sistêmico se dá devido às relações em cadeia existentes entre as empresas insolventes e outras empresas, como empresas de fornecedores, por exemplo. Além disso, funcionários, investidores e demais stakeholders também são afetados. Assim, torna-se relevante, portanto, antever situações de insolvência, para que as empresas possam evitar o encerramento das atividades devido à falência.

Diante disso, vários estudos de previsão foram desenvolvidos nos últimos anos, inicialmente envolvendo técnicas estatísticas tradicionais como análise discriminante (ALTMAN, 1968) e regressão logística (OLSHON, 1980). Posteriormente, técnicas de aprendizado passaram a ser usadas, como redes neurais artificiais e máquinas de vetores de suporte (PRADO *et al.*, 2016), com o intuito de obter resultados cada vez mais robustos. Existem, no entanto, diferentes técnicas de previsão, que podem ser classificadas em técnicas de caixa branca, caixa cinza e caixa preta, sendo geralmente assumido que as técnicas de caixa preta são as menos interpretáveis e mais robustas.

Apesar do desenvolvimento do tema, e diante das diferentes técnicas de previsão existentes, faz-se necessário o desenvolvimento de modelos atualizados, que comparem os diferentes tipos de técnicas existentes e que considerem analisar o poder preditivo de tais técnicas face a eventos não previstos, como a pandemia de COVID-19, que pode ser de difícil mensuração (KITOWSKI; KOWAL-PAWUL; LICHOTA, 2022).

Para tentar preencher essa lacuna, o objetivo deste estudo foi de comparar técnicas estatísticas (análise discriminante, regressão logística) e de aprendizado de máquina das três classificações principais: Caixa branca (k-vizinhos mais próximos e árvore de decisão), caixa cinza (floresta aleatória) e caixa preta (RNAs e SVMs), considerando o período de pandemia como amostra de teste para avaliar o desempenho dos modelos de previsão. Para isso, foram definidos os indicadores econômico-financeiros mais relevantes para a elaboração dos modelos, a amostra de teste foi definida considerando as empresas que entraram em situação de insolvência nos anos de 2020 e 2021, com as técnicas de previsão e os resultados das técnicas de previsão foram comparados.

A partir dos resultados, os indicadores econômico-financeiros mais relevantes e que foram considerados no estudo foram X_{24-w} , X_{28-w} , X_{15-w} , X_{39-w} , X_{40-w} e X_{14-w} . No que se refere à acurácia, a técnica com melhor percentual foi a de árvore de decisão (DT), seguida pela

de floresta aleatória (RF), regressão logística e análise discriminante, e, por fim, rede neural artificial (RNA), k-vizinhos mais próximos (KNN) e máquina de vetores de suporte (SVM). Tal resultado mostrou que as técnicas de caixa branca e caixa cinza tiveram desempenhos melhores de acurácia em relação às técnicas de caixa preta (RNA e SVM), resultados que estão de acordo com o estudo de Obermann e Waack (2015), estão parcialmente de acordo com o estudo de Barboza, Kimura e Altman (2017), e contrapõem estudos como os de Wilson e Sharda (1994), West (2000) e Baesens *et al.* (2003).

Sendo assim, considerando as empresas que entraram em situação de insolvência no período de pandemia como amostra de teste, os resultados apontaram para a superioridade de uma técnica de caixa branca (DT), seguida de uma de caixa cinza (RF) e das técnicas estatísticas tradicionais (LDA e LR). Diferentemente do que era esperado (OBERMANN; WAACK, 2015), as técnicas de caixa preta (RNA e SVM) apresentaram os piores desempenhos de acurácia.

Por tudo isso, considera-se que este estudo apresentou implicações acadêmicas, gerenciais e sociais. No que se refere às implicações acadêmicas, o presente estudo definiu entre uma grande variedade de indicadores econômico-financeiros, um número reduzido de variáveis para a composição de modelos de previsão. Adicionalmente, considerou-se o período de pandemia como amostra de teste, e verificou-se que diante desse evento não previsto, técnicas mais interpretáveis (DT e RF) e tradicionais (LDA e LR) apresentaram desempenhos superiores às técnicas de caixa preta (RNA e SVM), que geralmente são tratadas como mais robustas pela literatura (OBERMANN; WAACK, 2015).

Ao se tratar das implicações gerenciais, considera-se que este estudo proporciona informações que auxiliarão os tomadores de decisão a aderir estratégias mais assertivas, para garantir que a saúde financeira das empresas seja mantida. Assim, a antecipação de situações de insolvência por meio dos indicadores econômico-financeiros e de técnicas de previsão, medidas preventivas sejam tomadas para que as empresas não cheguem à situação extrema de falência.

Por fim, no que se refere às implicações sociais, menciona-se que a antecipação da situação de insolvência beneficia a sociedade como um todo. Isso se deve ao fato de que todos os agentes ligados à empresa que pode se tornar insolvente se beneficiarão com a antecipação de uma possível falência, uma vez que as empresas adotariam estratégias preventivas que contribuiriam para a sua continuidade e que beneficiará os investidores, fornecedores, funcionários, entre outros interessados na empresa.

Apesar disso, o presente estudo apresentou limitações. Entre elas, menciona-se que apenas empresas brasileiras não financeiras e com ações negociadas na B3 foram consideradas.

Devido às diferenças entre países desenvolvidos e emergentes, sugere-se que novos estudos considerando outros países sejam considerados, sendo eles outros países emergentes e também países desenvolvidos. Por meio de tais estudos, será possível comparar o desempenho de diferentes técnicas de previsão diante do evento não previsto da pandemia de COVID-19, sendo possível identificar se as técnicas mais interpretáveis e as tradicionais apresentarão melhores resultados se comparadas às de caixa preta em diferentes contextos institucionais.

Ademais, menciona-se também que foram considerados para a a previsão dados de apenas um ano antes de a empresa se encontrar em situação de insolvência. Para estudos futuros, sugere-se que métricas que refletem o aumento ou modificação de variáveis sejam consideradas, conforme Barboza, Kimura e Atman (2017), que afirmaram que as dificuldades que levarão à situações de insolvência podem ocorrer ao longo do tempo não somente um ano antes do evento de insolvência. Por meio de tal estudo, será possível uma análise mais aprofundada das técnicas de previsão

Por fim, menciona-se que apenas indicadores econômico-financeiros foram considerados para a elaboração dos modelos preditivos. Devido à existência de fatores institucionais, como a regulação, sugere-se que estudos futuros considerem também fatores do ambiente institucional que possam influenciar no risco de insolvência das empresas. Assim, a análise poderá se tornar mais completa, fato que pode proporcionar melhores desempenhos aos modelos de previsão de insolvência.

REFERÊNCIAS

- AKIAMA, S. R. **Probabilidade de Inadimplência de Grandes Empresas no Sistema Financeiro Nacional**. Dissertação (Mestrado em Ciências Contábeis) – Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 2008.
- ALTMAN, E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, n. 4, p. 589–609, 1968.
- ALTMAN, E.; BAIDYA, T.; DIAS, L. Previsão de problemas financeiros em empresas. **Revista de Administração de Empresas**, v. 19, n. 1, p. 17–28, 1979.
- AMBRÓSIO, P. E. **Redes Neurais Artificiais no Apoio ao Diagnóstico Diferencial de Lesões Intersticiais Pulmonares**. 2002. Dissertação (Mestrado em Física Aplicada à Medicina e Biologia) – Faculdade de Filosofia, Ciências e Letras, Universidade de São Paulo, Ribeirão Preto, 2002.
- AOKI, S.; HOSONUMA, Y. **Bankruptcy prediction using decision tree**. (H. Takayasu, Ed.) **Application of econophysics, proceedings**, 2004.
- BAESENS, B. *et al.* Benchmarking state-of-the-art classification algorithms for credit scoring. **Journal of the Operational Research Society**, v. 54, n. 6, p. 627–635, jun. 2003.
- BARBOSA, B. H. G. **Computação Evolucionária e Máquinas de Comitê na Identificação de Sistemas Não-Lineares**. Tese (Doutorado em Engenharia Elétrica) – Universidade Federal de Minas Gerais, Belo Horizonte, 2009.
- BARBOZA, F.; KIMURA, H.; ALTMAN, E. Machine learning models and bankruptcy prediction. **Expert Systems With Applications**, v. 83, p. 405–417, 2017.
- BATTISTON, S. *et al.* Credit chains and bankruptcy propagation in production networks. **Journal of Economic Dynamics and Control**, v. 31, n. 6, p. 2061–2084, 2007.
- BEAVER, W. H. Financial Ratios as Predictors of Failure. **Journal of Accounting Research**, v. 4, p. 71–111, 1966.
- BISHOP, C. M. **Neural networks and their applications**. Oxford: Clarendon Press, 1995.
- BRAGA, A.; LUDERMIR, T.; CARVALHO, A. **Redes Neurais Artificiais: Teoria e aplicações**. Rio de Janeiro: LTC Editora, 2000.
- BREIMAN, L. Random Forests. **Machine Learning**, v. 45, p. 5–32, 2001.
- BRITO, G.; ASSAF NETO, A.; CORRAR, L. Sistema de classificação de risco de crédito: uma aplicação a companhias abertas no Brasil. **Revista Contabilidade & Finanças**, v. 20, n. 51, p. 28–43, 2009.
- CHO, S.; HONG, H.; HA, B. C. A hybrid approach based on the combination of variable selection using decision trees and case-based reasoning using the Mahalanobis distance: For bankruptcy prediction. **Expert Systems With Applications**, v. 37, n. 4, p. 3482–3488, 2010.

CORRAR, L.; PAULO, E.; DIAS FILHO, J. **Análise multivariada para cursos de Administração, Ciências Contábeis e Economia**. São Paulo: Atlas, 2007.

CORTES, C.; VAPNIK, V. Support-Vector Networks. **Machine Learning**, v. 20, p. 273–297, 1995.

DIMITRAS, A. I. *et al.* Business failure prediction using rough sets. **European Journal of Operational Research**, v. 114, n. 2, p. 263–280, 1999.

ELIZABETSKY, R. **Um modelo matemático para decisão no banco comercial**. [s.l.] Universidade de São Paulo, 1976.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel, SPSS e Stata**. 1. ed. Rio de Janeiro: Elsevier, 2017.

FREIRE, M. *et al.* Aderência às normas internacionais de contabilidade pelas empresas brasileiras. **Revista de Contabilidade e Organizações**, v. 6, n. 15, p. 3–22, 2012.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5ª ed. Brasil: McGraw Hill Brasi, 2011.

GAVUROVA, B. *et al.* Analysis of Impact of Using the Trend Variables on Bankruptcy Prediction Models Performance. **Ekonomicky Casopis**, v. 65, n. 4, p. 370–383, 2017.

HAIR, J. *et al.* **Análise Multivariada de Dados**. 6ª edição ed. São Paulo: Bookman Companhia Editora, 2006.

HILLEGEIST, S. A. *et al.* Assessing the probability of bankruptcy. **Review of Accounting Studies**, v. 9, n. 1, p. 5–34, 2004.

HORTA, R. A. M.; ALVES, F. J. S.; CARVALHO, F. A. A. DE. Seleção de atributos na previsão de insolvência: aplicação e avaliação usando dados brasileiros recentes. **RAM. Revista de Administração Mackenzie**, v. 15, n. 1, p. 125–151, 2014.

HUANG, Z. *et al.* Credit rating analysis with support vector machines and neural networks: a market comparative study. **Decision Support Systems**, v. 37, n. 4, p. 543–558, 2004.

IZBICK, R.; SANTOS, T. **Aprendizado de máquina: uma abordagem estatística**. 1. ed. São Carlos: [s.n.].

KANAPRAM, D. T. *et al.* Graph-Powered Interpretable Machine Learning Models for Abnormality Detection in Ego-Things Network. **Sensors**, v. 22, n. 6, 2022.

KANITZ, S. Como Prever Falências de Empresas. **Revista Negócios em Exame**, p. 95–102, 1974.

KITOWSKI, J.; KOWAL-PAWUL, A.; LICHOTA, W. Identifying Symptoms of Bankruptcy Risk Based on Bankruptcy Prediction Models—A Case Study of Poland. **Sustainability**, v. 14, n. 3, 2022.

MATIAS, A. **Contribuição às técnicas de análise financeira: um modelo de concessão de crédito**. Monografia – Trabalho apresentado ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, São Paulo, 1978.

MILLER, T. Explanation in artificial intelligence: Insights from the social sciences. **Artificial Intelligence**, v. 267, n. July, p. 1–38, 2019.

MUELLER, A. .; GUIDO, S. **Introduction to Machine Learning with Python: A Guide for Data Scientists**. 1. ed. [s.l.] O'Reilly Media, 2019.

MUELLER, J. P.; MASSARON, L. **Machine Learning for dummies**. 2. ed. New Jersey: Hoboken, 2021.

OBERMANN, L.; WAACK, S. Demonstrating non-inferiority of easy interpretable methods for insolvency prediction. **Expert Systems With Applications**, v. 42, n. 23, p. 9117–9128, 2015.

ODOM, M. D.; SHARDA, R. **A Neural Network Model For Bankruptcy Prediction**. Proceedings of the IEEE International Conference on Neural Network, 2, 163–168. **Anais...**1990

OHLSON, J. A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. **Journal of Accounting Research**, v. 18, n. 1, p. 109–131, 1980.

PEDRO JÚNIOR, V. **Modelo de previsão de insolvência para Grandes empresas: uma abordagem com GEE**. 2011. Dissertação (Mestrado profissional em Economia) - Insper Instituto de Ensino e Pesquisa, São Paulo, 2011.

PINTO, A. C. O poder preditivo dos modelos com aprendizado de máquina é superior aos modelos tradicionais para análise do risco de crédito? **Revista Debates em Economia Aplicada – REDEA**, v. 1, n. 1, 2021.

PRADO, J. W. **Determinantes e implicações da estrutura de capital, da estrutura de propriedade e da governança corporativa: um modelo multiteórico de análise**. 2019. Tese (Doutorado em Administração) – Universidade Federal de Lavras, Lavras, 2019.

PRADO, J. *et al.* Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). **Scientometrics, Budapeste**, v. 106, p. 1007–1029, 2016.

PRADO, J. **Risco de crédito: uma abordagem utilizando análise discriminante, regressão logística e redes neurais artificiais**. 2016. Dissertação (Mestrado em Administração) – Universidade Federal de Lavras, Lavras, 2016.

PRADO, J. *et al.* Previsão da insolvência empresarial utilizando redes neurais artificiais. **Revista Gestão e Desenvolvimento, Novo Amburgo**, v. 17, n. 2, p. 136–162, 2020.

QUINLAN, J. R. Induction of decision trees. **Machine Learning 1:**, v. 1, p. 81–106, 1986.

SANTOS, A. S. **Previsão de insolvência corporativa: uma análise de empresas brasileiras de capital aberto por meio de aprendizado de máquina.** [s.l.] Universidade Federal da Paraíba, 2021.

SANVICENTE, A. Z.; MINARDI, A. M. A. F. Identificação de indicadores contábeis significativos para previsão de concordata de empresas. **Instituto Brasileiro de Mercado de Capitais, WorkingPaper**, São Paulo, p. 1-12, 1998.

SICSÚ, A. **Credit scoring: desenvolvimento, implantação, acompanhamento.** São Paulo: Blucher, 2010.

SILVA, A.; FONSECA, E.; LITTIG, R. Modelo Dinâmico Fleuret: Uma Análise Das Empresas Com Pedidos De Recuperação Judicial Listadas Na Bovespa. **Revista Científica Multidisciplinar Núcleo do Conhecimento**, v. 2, n. 5, p. 114–137, 2019.

SILVA, J. (1982). **Modelos para classificação de empresas com vistas a concessão de crédito.** Dissertação (Mestrado em Administração) - Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas. São Paulo, 1982.

SJÖBERG, J. *et al.* Nonlinear Black-Box Modeling in System Identification: a Unified Overview. **Automatica**, v. 31, n. 12, p. 1691–1724, 1995.

SOARES, R.; REBOUÇAS, S. Avaliação do desempenho de técnicas de classificação aplicadas à previsão de insolvência de empresas de capital aberto brasileiras. **Revista ADM. Made**, v. 18, n. 3, p. 40–61, 2014.

STÜPP, D. R. **Previsão de insolvência a partir de indicadores contábeis: evidências de empresas listadas na BM&FBOVESPA nos anos 2004-2013.** Dissertação (Mestrado em Contabilidade) - Universidade Federal de Santa Catarina, Florianópolis, 2015.

SUMIHARA FILHO, H.; SLEEGERS, L. C. Valores de referência para os principais indicadores de acurácia dos modelos de escoragem. **Serasa Experian - Tecnologia de Crédito**, São Paulo, n. 73, p. 31-45, ago. 2010.

TELES, G. *et al.* Comparative study of support vector machines and random forests machine learning algorithms on credit operation. **Software - Practice and Experience**, v. 51, n. 12, p. 2492–2500, 2021.

WELESZCZUK, J.; KOSIŃSKA-SELBI, B.; CHOLEWIŃSKA, P. Prediction of Polish Holstein's economical index and calving interval using machine learning. **Livestock Science**, v. 264, p. 105039, 2022.

WEST, D. Neural network credit scoring models. **Computers & Operations Research**, v. 27, p. 1131–1152, 2000.

WILSON, R. L.; SHARDA, R. Bankruptcy prediction using neural networks. **Decision Support Systems**, v. 11, p. 545–557, 1994.

ARTIGO 3 – EFEITOS DA ESTRUTURA DE CAPITAL NO RISCO DE INSOLVÊNCIA

RESUMO

O objetivo deste estudo foi analisar a influência da estrutura de capital no risco de insolvência de empresas brasileiras não financeiras com ações negociadas na B3, analisando também os efeitos da pandemia na relação entre estrutura de capital e risco de insolvência. Para o desenvolvimento do estudo, foi usada a regressão multinível de dados em painel, com dados trimestrais de 2010 a 2021. Os resultados mostraram que as variáveis de endividamento apresentaram efeitos significativos positivos, corroborando o proposto pela teoria do *trade-off*. A variável endividamento financeiro e seu termo quadrático apresentaram resultados significativos com coeficientes negativo e positivo, respectivamente, confirmando a existência de relação quadrática (em U) entre o endividamento financeiro e o risco de insolvência, devido ao benefício fiscal, que compensa, até certo ponto, os custos de falência associados ao endividamento. Em relação ao efeito moderador da pandemia, apenas a hipótese de que a pandemia de COVID-19 fortalece a influência positiva do endividamento de longo prazo no risco de insolvência foi confirmada. O risco de insolvência se constitui como um importante fator para diversos agentes da sociedade. Os estudos sobre insolvência foram desenvolvidos principalmente no sentido de elaboração de modelos de previsão, mas a identificação de fatores determinantes do risco de insolvência contribui para entender como esse problema surge e auxilia na sua prevenção. A estrutura de capital influencia no risco de insolvência, no entanto, visões distintas sobre os efeitos da estrutura de capital no valor e no risco da empresa são encontradas.

Palavras-chave: Endividamento. Risco de falência. COVID-19.

ABSTRACT

The aim of this study was to analyze the influence of capital structure on the risk of insolvency of Brazilian non-financial companies with shares traded on B3, also analyzing the effects of the pandemic on the relationship between capital structure and risk of insolvency. For the development of the study, multilevel regression of panel data was used, with quarterly data from 2010 to 2021. The results showed that the debt variables had significant positive effects, corroborating the proposed by the trade-off theory. The financial indebtedness variable and its quadratic term showed significant results with negative and positive coefficients, respectively, confirming the existence of a quadratic (U-shaped) relationship between financial indebtedness and the risk of insolvency, due to the tax benefit, which compensates, to a certain extent, the bankruptcy costs associated with debt. Regarding the moderating effect of the pandemic, only the hypothesis that the COVID-19 pandemic strengthens the positive influence of long-term debt on the risk of insolvency was confirmed. The risk of insolvency is an important factor for various agents in society. Studies on insolvency were developed mainly in order to develop forecast models, but the identification of factors that determine the risk of insolvency contributes to understanding how this problem arises and helps in its prevention. The capital structure influences the risk of insolvency, however, different views on the effects of the capital structure on the value and risk of the company are found.

Keywords: Indebtedness. Bankruptcy risk. COVID-19.

1 INTRODUÇÃO

As empresas com insuficiência de saldo (patrimônio líquido negativo) são consideradas como insolventes (SILVA, 2010). No trabalho de Merton (1974), o autor introduziu a ideia de que o processo de falência é endógeno, estando atrelado à estrutura de capital. Conforme o estudo, as empresas são consideradas insolventes quando o valor de seus ativos for menor que o valor da dívida conforme o vencimento (CHAIA, 2003). Essa ideia é considerada como similar à de insolvência de saldo, sendo necessário analisar os efeitos da estrutura de capital.

Os estudos de previsão de insolvência foram desenvolvidos para a geração de modelos preditivos para a antecipação de problemas financeiros (ALTMAN, 1968; PRADO *et al.*, 2016; SHI; LI, 2019b). Usando principalmente indicadores econômico-financeiros, esse campo de estudos é considerado relacionado com a literatura de estrutura de capital. Isso ocorre porque a definição do nível de endividamento e a antecipação de problemas financeiros são considerados cruciais para manter a saúde financeira das empresas.

Diferentes fatores influenciam no risco de insolvência, entre eles a estrutura de capital. Uma vez que a estrutura de capital indica a forma de financiamento da empresa e, portanto, o seu nível de endividamento, ela é considerada um dos fatores determinantes de tal risco. Sendo assim, uma estrutura de capital inadequada influencia no risco de insolvência e no desempenho da empresa (TAO *et al.*, 2020).

A estrutura de capital é entendida como a estrutura de financiamento/endividamento de longo prazo usada pelas empresas. As empresas são financiadas por capital próprio ou por capital de terceiros. Sendo assim, a estrutura de capital mostra as fontes de recursos usadas para financiar os ativos (BITTENCOURT; ALBUQUERQUE, 2018; MYERS, 1984). A literatura existente aponta para a existência de diferentes teorias sobre a estrutura de capital. Entre elas são encontrados defensores de que a estrutura de capital seja irrelevante para o valor da empresa (MODIGLIANI; MILLER, 1958), e defensores de que a estrutura de capital seja relevante para o valor da empresa (DURAND, 1952).

Entre as teorias de estrutura de capital destaca-se a teoria do *trade-off*, também conhecida como teoria dos custos de falência (PRADO, 2019). De acordo com essa teoria, a estrutura de capital é um fator relevante, existindo um nível ótimo entre o capital próprio e o de terceiros, no qual há um equilíbrio entre os custos de falência e benefícios fiscais provenientes do uso de dívidas (capital de terceiros). Isso ocorre porque há incidência de juros no capital de terceiros, que diminuem o lucro tributável, conseqüentemente diminuindo o valor dos impostos pagos, o que garante um benefício fiscal (MYERS, 1984; PRADO, 2019).

Outro fator relevante que impactou a economia e que pode ter afetado o efeito da estrutura de capital no risco de insolvência foi a pandemia de COVID-19. De acordo com a CNN Brasil (2021), uma pesquisa realizada pelo Economática ao Estadão apontou que entre as 218 empresas não financeiras de capital aberto analisadas, houve uma redução de 81,9% no lucro líquido do primeiro semestre de 2020, se comparado ao mesmo período de 2019, com uma queda de ganhos de R\$ 49,2 bilhões para R\$ 8,9 bilhões. Esses resultados evidenciam os efeitos da pandemia na economia, especificamente nas empresas brasileiras não financeiras de capital aberto, que são o foco deste estudo. Sendo assim, analisou-se o efeito moderador da pandemia de COVID-19 sobre o impacto da estrutura de capital no risco de falência.

Devido à relevância do tema, a identificação de fatores determinantes do risco de insolvência contribuem para entender como esse problema surge e auxilia na sua prevenção. A estrutura de capital pode influenciar no risco de insolvência, uma vez que o nível de endividamento pode ser um indício de problemas financeiros. No entanto, encontram-se na literatura visões distintas sobre os efeitos da estrutura de capital no valor e no risco da empresa. Mesmo acreditando na existência dos custos de falência, Miller (1977) destacou que estes custos são considerados desproporcionais ao benefício fiscal, visão oposta à defendida pela teoria do *trade-off*. Ademais o endividamento e risco de insolvência são menos analisados conjuntamente e existem resultados controversos sobre a relação entre o endividamento e o risco de insolvência (LOTT, 2019).

Sendo assim, o presente estudo visa aprofundar o conhecimento analisando o impacto da estrutura de capital no risco de insolvência. Neste contexto, o presente estudo tem como **objetivo** analisar a influência da estrutura de capital no risco de insolvência de empresas brasileiras não financeiras de capital aberto negociadas na B3.

Para a elaboração do estudo foram desenvolvidos os seguintes passos: analisar o risco de insolvência de empresas não financeiras negociadas na B3; analisar o efeito da estrutura de capital no nível de risco de insolvência de empresas não financeiras negociadas na B3; analisar os efeitos da pandemia de Covid-19 na relação entre estrutura de capital e nível de risco de insolvência. Para as análises foram usados dados presentes nos relatórios financeiros dessas empresas. Esses dados foram coletados através do banco de dados Economática.

Adicionalmente, a teoria de estrutura de capital postula sobre o risco de altos níveis de endividamento, ou seja, o financiamento por capital de terceiros é benéfico, no entanto o uso exclusivo de capital de terceiros não é recomendado (MODIGLIANI; MILLER, 1963). Segundo Lott, Tenenwurcel e Camargos (2021), ainda não existe uma definição teórica objetiva em finanças sobre como os tomadores de decisão definem o grau de endividamento da empresa.

Ademais, os autores mencionam que na literatura sobre estrutura de capital também não há um consenso.

É importante ressaltar também que as teorias sobre estrutura de capital foram amplamente pesquisadas considerando países desenvolvidos (OZ; SIMGA-MUGAN, 2018). No entanto, apesar da importância das teorias desenvolvidas, é importante mencionar que os países emergentes possuem particularidades consideradas relevantes para definição da estrutura de capital, bem como para o risco de insolvência. Isso ocorre devido ao fato de que diferentes países possuem contextos institucionais específicos, devido a fatores como a formação histórica.

Os países da América Latina, especificamente, possuem características que os diferem de países de outras localidades, como incerteza política, violência, reformas e reversões pró-mercado e isolamento geográfico (CUERVO-CAZURRA, 2016). Essas diferenças fazem com que sejam necessárias análises específicas desses países, uma vez que o conhecimento produzido em diferentes contextos, como os de países Asiáticos ou países desenvolvidos, não devem ser generalizados para esses países.

Diante disso, a análise de economias emergentes, como o Brasil, se faz relevante para a ampliação do conhecimento. Especificamente, neste estudo estendemos o conhecimento sobre estrutura de capital e risco de insolvência, considerando o contexto único de um país da América Latina. O Brasil foi escolhido por ser uma importante economia emergente da América Latina, tratando-se do país com maior percentual do PIB da região, que de acordo com dados da CEPAL (2021), corresponde a 33,25% do PIB total da região. Ademais, foi escolhido devido à disponibilidade de dados necessários para a realização das análises propostas no estudo.

A estrutura do trabalho foi definida da seguinte forma: nos próximos tópicos estiveram o referencial teórico do estudo, abrangendo a Teoria do *Trade-off*, estudos sobre estrutura de capital e as hipóteses testadas, a metodologia do estudo, com o detalhamento dos procedimentos de coleta de dados, amostra, variáveis e procedimentos de análises, resultados e discussão do estudo e conclusões.

2 REFERENCIAL TEÓRICO

Neste tópico encontra-se o referencial teórico do artigo. Nele foram apresentados a teoria do *Trade-off*, estudos envolvendo a estrutura de capital e as hipóteses testadas.

2.1 Teoria do *Trade-off*

A teoria do *trade-off* considera basicamente os benefícios advindos de benefícios fiscais e os custos de falência. Isso significa que a teoria reconhece que o capital de terceiros proporciona um benefício fiscal. No entanto, deve haver um equilíbrio entre o capital de terceiros e capital próprio, uma vez que o custo da dívida pode ocasionar dificuldades financeiras e falência (RICCA; JUCÁ; HADAD JUNIOR, 2021). Logo, o ideal é que a empresa mantenha um equilíbrio, substituindo capital próprio por capital de terceiros e capital de terceiros por capital próprio quando necessário, para minimizar o custo de capital e maximizar o valor da empresa (MYERS, 1984; RICCA; JUCÁ; HADAD JUNIOR, 2021).

A teoria do *trade-off* possui duas visões sobre o nível ótimo de endividamento. Na teoria do *trade-off* estático tem-se a ideia de que existe um único ponto-ótimo estabelecido, que deve ser alcançado pelas empresas. Na teoria do *trade-off* dinâmico, por sua vez, o entende-se que o nível ótimo de endividamento depende também de fatores do ambiente externo, podendo variar ao longo do tempo (LOTT, 2019).

Sendo assim, apesar de a ideia inicial de haver um equilíbrio que maximize o valor da empresa, pode haver um desvio da meta de alavancagem por parte das empresas (RICCA; JUCÁ; HADAD JUNIOR, 2021). A partir de tal desvio, a estrutura de capital estática pode dar espaço para uma estrutura de capital dinâmica. Na estrutura de capital dinâmica as empresas usam a emissão de dívidas para financiar investimentos em certo momento, ação que diminui a capacidade futura de endividamento, como melhores condições proporcionadas por credores (DEANGELO; DEANGELO; WHITED, 2011; RICCA; JUCÁ; HADAD JUNIOR, 2021).

De maneira complementar grande parte dos modelos que abordaram a estrutura de capital considera o nível de endividamento como uma decisão estática, mas no mundo real, o nível de endividamento é ajustado pelas empresas dependendo das mudanças que influenciam seu valor (NENU; VINTILĂ; GHERGHINA, 2018).

2.2 Estudos envolvendo a estrutura de capital

Nenu, Vintilă e Gherghina (2018) desenvolveram seu estudo para analisar os fatores de impacto na alavancagem, lucratividade e no risco de empresas listadas na Bolsa de valores de Bucareste. De acordo com os autores, os estudos iniciais sobre a teoria de estrutura de capital foram desenvolvidos considerando condições de mercado perfeito, apresentando proposições bastante teóricas. Para eles, mesmo que a teoria de estrutura de capital tenha se desenvolvido, ela continua sendo um desafio, e em mercados emergentes existem particularidades que devem ser consideradas.

Por meio do estudo, os autores encontraram uma relação positiva entre a alavancagem financeira e o tamanho da empresa (consistente com a teoria do *trade-off*). No entanto, identificaram que a estrutura de capital impactou o desempenho das empresas de maneira diferente. Por fim, menciona-se que os autores identificaram uma relação negativa entre rentabilidade e endividamento de curto e longo prazo.

No trabalho de Hovakimian, Kayhan e Titman (2012) os autores testaram a teoria do *trade-off* estático, analisando a probabilidade de inadimplência, que desempenha um papel central na teoria. A partir do estudo, os resultados obtidos foram contrários às previsões da teoria. Ou seja, os resultados do estudo indicaram que empresas com maiores custos de falência (que são menores em tamanho e com menor tangibilidade de ativos) tendem a apresentar estruturas de capital com risco de falência maior.

No estudo de Ricca, Jucá e Hadad Junior (2021), os autores tiveram como objetivo analisar o nível de conservadorismo de empresas em relação à estrutura de capital, procurando entender por que as empresas não usam integralmente o benefício fiscal da dívida. A partir do estudo, as evidências mostraram que o nível de alavancagem financeira das empresas brasileiras não apresentou um nível ótimo, mesmo que os benefícios da dívida superassem os custos de falência.

Entre os motivos que podem ter levado as empresas brasileiras a apresentarem uma posição conservadora em relação à alavancagem, os autores citam fatores do ambiente macroeconômico e crises institucionais no país. A Lei de falências também foi citada como um fator que pode ter impactado, fato relacionado à insegurança de sua aplicação e na manutenção de um equilíbrio entre o direito dos credores e dos devedores. Os resultados sugeriram também que aparentemente as empresas usam outros tipos de benefícios fiscais.

Face à importância de estudos de estrutura de capital em países emergentes, Haron *et al.* (2021) realizaram um estudo considerando empresas da Indonésia. No estudo de Haron *et*

al. (2021), foram analisados fatores no nível da firma, no nível da indústria e o efeito do nível de concentração de propriedade nas decisões sobre estrutura de capital dinâmica, analisando também as teorias de governança corporativa. Como resultado, os autores concluíram que, consistente com a teoria do *trade-off* empresas que estão crescendo rapidamente e empresas que operam em indústrias altamente concentradas possuem estruturas de capital mais alavancadas, se beneficiando do escudo fiscal gerado pelo uso de capital de terceiros.

Em contrapartida, as empresas com mais tempo de mercado e com liquidez, lucro e ativos tangíveis e intangíveis altos e que atuam em mercados dinâmicos, apresentam um menor nível de endividamento, para evitar o risco de falência e suas consequências. Sendo assim, estas últimas dão preferência ao financiamento hierárquico (teoria do *pecking order*), reduzindo sua dependência por dívidas.

A literatura aponta também que fatores externos impactam o nível de endividamento das empresas, bem como no risco de insolvência (HACKBARTH; MIAO; MORELLEC, 2006; TINOCO; WILSON, 2013). No estudo de Jõeveer (2013), foi analisada a importância de fatores macroeconômicos, institucionais e da empresa como determinantes da estrutura de capital de empresas de nove países do leste europeu de 1995 a 2002.

Como resultado, verificou-se que para empresas de capital aberto, os fatores do nível da empresa foram os mais relevantes, enquanto para as empresas não listadas em bolsa e de menor tamanho os fatores específicos do país foram os principais determinantes da alavancagem. De maneira geral, os resultados do estudo indicaram que as variáveis do país foram determinantes significativos do nível de alavancagem, em especial para as empresas não listadas.

Ao considerar a teoria do *trade-off*, Jõeveer (2013) afirma que essa teoria prevê que empresas de tamanho maior, com mais ativos tangíveis e mais lucrativas possuem mais acesso aos benefícios fiscais, tendendo a serem mais alavancadas. De maneira complementar, o autor afirma que fatores do ambiente macroeconômico também influenciam no nível de alavancagem das empresas, como os impostos locais e a legislação de falência. Isso é explicado pelo fato de que em localidades com taxas de impostos maiores, o benefício fiscal dos juros tenderá a ser maior, induzindo a uma maior alavancagem. A inflação é outro fator previsto como positivamente relacionada à alavancagem, pois o valor real das deduções fiscais sobre a dívida é maior. Em relação à taxa de juros (custo da dívida), espera-se uma relação negativa com a alavancagem.

No trabalho de Tinoco e Wilson (2013), os autores analisaram o uso combinado de variáveis contábeis, variáveis de mercado e macroeconômicas como determinantes do risco de crédito corporativo. Para o estudo, os autores analisaram uma amostra contendo 23.218

observações (empresa-ano) durante o período de 1980 a 2011. Os autores identificaram que o uso combinado das variáveis apresentou resultados satisfatórios, sendo útil combinar dados de mercado, dados contábeis e macroeconômicos, diferentemente do que comumente é feito. Os autores ressaltaram que geralmente tais dados são usados separadamente e com uma visão de que são concorrentes não complementares.

Corroborando o trabalho de Tinoco e Wilson (2013), Carvalho, Curto e Primor (2022) analisaram a relação entre variáveis macroeconômicas e o risco de inadimplência de empresas da Zona do Euro. Por meio do estudo, os resultados mostraram que as variáveis macroeconômicas impactaram na inadimplência de crédito das empresas e proporcionaram um aumento significativo na precisão dos modelos preditivos de risco de crédito. O estudo proporcionou uma análise ampla, que também mostrou que a relação das variáveis macroeconômicas com o risco de crédito variou entre os diferentes países da Zona do Euro. Entre as variáveis destacadas estiveram o PIB (produto Interno Bruto) e a inflação, que de acordo com os autores influenciaram negativamente o risco de inadimplência, e a taxa de juros, que influenciou positivamente o risco de inadimplência.

No trabalho de Jang, Rhee e Yoon (2016), o objetivo foi desenvolver um modelo estrutural que incorporasse riscos de ciclos de negócios (do ambiente macroeconômico) e *jump risks* (risco de salto: associado à volatilidade no valor de mercado), para explicação das taxas de inadimplência. Por meio do estudo, o modelo pode prever probabilidades reais de inadimplência e superar subestimações de riscos de crédito principalmente para empresas com classificações de crédito altas, que de acordo com os autores tem sido uma limitação de modelos estruturais atuais. As variáveis macroeconômicas foram relevantes nos modelos de risco de crédito, e os *spreads* de rendimento e taxas de inadimplência podem ser dependentes de fatores do ambiente econômico.

De maneira complementar, Hackbarth, Miao e Morellec (2006) analisaram o efeito do ambiente macroeconômico no risco de crédito e na estrutura de capital dinâmica. A partir do estudo, os autores identificaram que os regimes de ciclos de negócios (fatores macroeconômicos) foram relevantes para o risco de crédito e para a determinação da estrutura de capital. Apesar da importância dos ciclos de negócios, os autores ressaltaram que até então poucos trabalhos haviam dado atenção ao ritmo e às mudanças na estrutura de capital entre regimes de ciclos de negócios.

2.3 Estrutura de capital em mercados emergentes

De acordo com Duran e Stephen (2020), a política relacionada à definição da estrutura de capital de empresas de países em desenvolvimento e, especificamente, de países da América Latina pode ser considerada como única. Isso se deve ao fato de que os ambientes institucionais desses países possuem características diferentes de países desenvolvidos, como, por exemplo, ambientes institucionais instáveis e mercados de capitais, instituições financeiras e regulação menos desenvolvidos. Esses fatores impactam na estrutura de capital ótima, uma vez que a exploração do uso de capital de terceiros pode ser influenciado pelas características institucionais desses países (BASTOS; NAKAMURA; BASSO, 2009).

De maneira complementar, segundo Brito, Corrar e Batistella (2007), o Brasil possui características específicas, que fazem com que as evidências empíricas de países desenvolvidos relacionadas aos determinantes do nível de endividamento sejam diferentes das evidências encontradas em estudos brasileiros, que apresentam também resultados mistos. De acordo com os mesmos autores, o Brasil possui características como taxas de juros altas, mercado de capitais restrito, forte restrição no que se refere às fontes de financiamento por meio de capital de terceiros de longo prazo e alta concentração de controle acionário nas empresas. Esses fatores podem ser entendidos como ineficiências do país que também influenciam a determinação da estrutura de capitais das empresas.

Especialmente as altas taxas de juros fazem com que as empresas tenham baixos níveis de endividamento, devido ao alto custo de financiamento decorrente de tais taxas de juros (BRITO; CORRAR; BATISTELLA, 2007). Ademais, Brito, Corrar e Batistella (2007) afirmam que devido à existência de diferentes naturezas das fontes de financiamento, o custo de capital, além de ser influenciado pelo risco do tomador, também sofre influência da natureza da fonte de financiamento, uma vez que existem linhas de crédito para tipos específicos para investimentos de longo prazo, que apresentam um custo financeiro inferior ao das linhas de crédito de curto prazo.

2.4 Estrutura de capital e risco de insolvência em períodos de crise

De acordo com Zeitun, Temimi e Mimouni (2017), a existência de crises econômicas se constitui como um aspecto desafiador para teorias financeiras, como, por exemplo, para as teorias de estrutura de capital. Segundo os autores, em períodos de crise os recursos destinados aos empréstimos tendem a ser reduzidos, mas ao mesmo tempo as empresas podem precisar

fazer uso de dívida para passar por esses períodos de crise, períodos estes que dificultam a geração de fluxo de caixa. Sendo assim, esses fatores podem fazer com que a estrutura de capital das empresas seja impactada, tanto pela escassez de financiamento disponíveis, quanto pela necessidade do uso de capital de terceiros para manter-se em funcionamento no período de crise.

Diante disso, em períodos de crises econômicas, as empresas tendem precisar modificar suas estratégias, fazendo-se necessário considerar tais períodos em novas análises (SILVA *et al.*, 2016). Um exemplo de crise econômica, de acordo com Silva *et al.* (2016), foi a crise das hipotecas *subprime*, iniciada nos Estados Unidos. Ela desencadeou uma crise mundial, impactando decisões referentes aos investimentos e também decisões de financiamento. Outro exemplo de crise econômica mundial foi a decorrente da pandemia de COVID-19, que causou impactos negativos na economia global, que podem demorar a ser superados, especialmente em países emergentes e em desenvolvimento (THE WORLD BANK, 2021, 2022).

Existem evidências de que a estrutura de capital das empresas se modificam após períodos de crises econômicas (ALVES; FRANCISCO, 2015; DURAN; STEPHEN, 2020; LYUBOV; HESHMATI, 2019). No estudo de Duran e Stephen (2020), os autores identificaram que o endividamento das empresas foi menor no período antes da crise financeira para empresas domésticas e multinacionais. Após a crise, os autores identificaram que as empresas multinacionais apresentaram maior nível de endividamento. De acordo com Duran e Stephen (2020), algumas possíveis explicações para o aumento do uso de capital de terceiros pelas empresas multinacionais da América Latina após a crise foi devido à existência de taxas de juros mais atrativas que estimularam as economias, assim como o acesso aos mercados de capitais globais.

Corroborando os resultados de Duran e Stephen (2020), Lyubov e Heshmati (2019) identificaram que a crise de 2008 influenciou positivamente o nível de endividamento de empresas coreanas listadas. No entanto, considerando a crise financeira asiática de 1997, os autores identificaram o endividamento médio diminuiu significativamente após essa crise.

De maneira complementar, Alves e Francisco (2015) analisaram os efeitos de fatores institucionais durante três crises financeiras (bolha, crise do *subprime* e crise da dívida soberana europeia) no endividamento de empresas de diversos países de todo o mundo. Como resultado, os autores identificaram o endividamento aumentou nos períodos de crise, sendo que o endividamento de longo prazo foi substituído pelo endividamento de curto prazo, o que implica em dificuldades financeiras, face ao aumento do risco de rolagem, em que as empresas fazem uso de uma nova dívida para pagar a dívida anterior.

Existem também, no entanto, evidências de que a crise financeira de 2008 diminuiu o endividamento de empresas (IWAKI, 2019; ZEITUN; TEMIMI; MIMOUNI, 2017). Especificamente, Iwaki (2019) identificou um efeito negativo da crise de 2008 na estrutura de capital de empresas dependentes essencialmente de bancos, ou seja, que não possuem acesso ao mercado de dívida pública.

De maneira semelhante, Zeitun, Temimi e Mimouni (2017) identificaram que a crise de 2008 influenciou negativamente o endividamento, fato também associado à menor oferta por parte dos credores. Ademais, Zeitun, Temimi e Mimouni (2017) descobriram que antes da crise, o principal determinante do endividamento das empresas era a demanda por capital de terceiros pelas empresas, e após a crise, além da demanda, a oferta de dívida pelos credores também passou a ser um determinante do nível de endividamento.

Além da estrutura de capital, em períodos de crise o risco de falência das empresas também pode ser influenciado. Isso ocorre pois em tempos de recessão as empresas tendem a ter seu desempenho econômico-financeiro comprometido, fato que faz com que o aumento do risco de insolvência se torne comum nesses períodos (CHAIA, 2003). Assim, uma vez que geralmente a oferta de crédito se torna menor em períodos de crise e algumas empresas necessitam de crédito para continuar com suas operações (ZEITUN; TEMIMI; MIMOUNI, 2017), tais empresas podem optar por dívidas de curto prazo, que podem aumentar o risco de rolagem e, conseqüentemente, o risco de insolvência (ALVES; FRANCISCO, 2015).

2.5 Desenvolvimento de hipóteses sobre a influência da estrutura de capital no risco de falência

Conforme já visto, a decisão sobre a estrutura de capital de uma empresa é considerada como complexa, podendo influenciar na saúde financeira das organizações (BRITO; CORRAR; BATISTELLA, 2007; NENU; VINTILĂ; GHERGHINA, 2018). Tal impacto pode ser explicado pela teoria do *trade-off*, que defende a existência de um benefício fiscal proveniente do uso de capital de terceiros, existindo, porém, a necessidade de atenção aos custos de falência, que podem se tornar mais altos do que o recomendado caso o endividamento se eleve, diminuindo a eficácia do benefício fiscal (MYERS, 2001). Isso ocorre porque o valor dos juros sobre o capital de terceiros é deduzido do valor no qual há a incidência de imposto de renda, diminuindo o valor de impostos pagos.

Ademais, deve-se destacar que o benefício fiscal proposto pela teoria do *trade off* advém basicamente das despesas financeiras incidentes sobre o capital de terceiros, no entanto, nem

todos os componentes do passivo geram o pagamento de juros. De acordo com Machado, Medeiros e Eid Júnior (2010), o passivo se divide em passivo oneroso e passivo não-oneroso.

O passivo oneroso corresponde à parcela do passivo em que há incidência de juros, ou seja, que gera despesas financeiras que proporcionam o benefício fiscal pelo uso do capital de terceiros. Já o passivo não-oneroso decorre basicamente das atividades operacionais da empresa, ou seja, como o próprio nome sugere, sobre esse tipo de passivo não há a incidência de despesas financeiras. Sendo assim, torna-se relevante levar em consideração essa diferença ao analisar os efeitos do endividamento no nível de risco de insolvência das empresas brasileiras não financeiras negociadas na B3, uma vez que espera-se que o benefício fiscal gerado pelo uso do capital de terceiros seja proveniente basicamente do passivo oneroso.

Apesar de tal benefício fiscal, as empresas devem estar atentas para que seu nível de endividamento não alcance patamares elevados. Essa preocupação é necessária, pois os custos de falência podem aumentar a tal ponto que o benefício fiscal do endividamento se torne insuficiente (MYERS, 2001). Sendo assim, as empresas devem procurar um equilíbrio entre o uso de capital próprio e capital de terceiros, para que tenham uma estrutura que minimize os custos (de capital e de falência), ou seja, uma estrutura de capital eficiente.

Face ao possível efeito do endividamento nos custos de falência, o aumento o risco de falência pode ser um efeito decorrente de altos níveis de endividamento. Conforme já mencionado e de acordo com a teoria do *trade-off*, caso o endividamento supere o nível ideal, os custos de falência podem superar o benefício fiscal, que é decorrente apenas do chamado passivo oneroso. Isso faz com que a empresa passe a ter mais risco de entrar em situação de insolvência. Face a tais suposições, espera-se que aumentos no nível de endividamento das empresas em geral influenciem positivamente no risco de insolvência, uma vez que o benefício fiscal não é proveniente de qualquer tipo de capital de terceiros, mas apenas daqueles que geram o pagamento de juros. Diante disso, as primeiras hipóteses do estudo foram propostas:

Hipótese 1a: A dívida líquida de curto prazo influencia positivamente o risco de insolvência

Hipótese 1b: O endividamento de curto prazo influencia positivamente o risco de insolvência.

Hipótese 1c: O endividamento de longo prazo influencia positivamente o risco de insolvência.

Conforme já mencionado, os juros incidem apenas sobre parte do passivo, que é o chamado passivo oneroso (MACHADO; MEDEIROS; EID JÚNIOR, 2010). Diante disso, espera-se que o benefício fiscal, defendido pela teoria do *trade off*, seja gerado apenas por parte do passivo. Ou seja, espera-se que a dedução do imposto de renda seja proveniente dos juros que incidem sobre o passivo oneroso, gerando o benefício fiscal para a empresa.

No entanto, de acordo com a teoria do *trade off*, quando o uso de capital de terceiros é demasiado, tal uso faz com que sejam gerados custos de falência que superarão os efeitos do benefício fiscal. Sendo assim, face ao benefício fiscal proveniente do passivo oneroso, espera-se que exista uma relação quadrática (em forma de U) entre o endividamento proveniente do passivo oneroso (dívida financeira bruta total) e o risco de insolvência. Assim, espera-se que o benefício fiscal dos juros diminua o risco de insolvência até certo ponto (nível ótimo de endividamento), a partir do qual os custos de falência superarão o benefício fiscal, aumentando o risco de falência da organização. Diante disso, a hipótese 1d do estudo foi levantada:

Hipótese 1d: O endividamento financeiro influencia negativamente e até certo ponto o risco de insolvência, após o qual a relação torna-se positiva.

Adicionalmente, diante da situação de pandemia, espera-se que o efeito da estrutura de capital no risco de insolvência tenha se modificado, sendo que a pandemia de COVID-19 pode ter apresentado um efeito moderador nessa relação. Conforme citado anteriormente, existem evidências de que fatores externos influenciam no risco de insolvência das empresas (CARVALHO; CURTO; PRIMOR, 2022; TINOCO; WILSON, 2013). A pandemia de COVID-19 pode ser considerada como um fator adverso do ambiente externo que causou grandes impactos na economia mundial, afetando empresas e a economia de todo o mundo (THE WORLD BANK, 2022).

Diante dos efeitos da pandemia de COVID-19 na economia mundial, bem como da paralização de diversas empresas devido ao isolamento social (FAHLENBRACH; RAGETH; STULZ, 2021), espera-se que as empresas tenham necessitado de recursos de terceiros para manter seus pagamentos em dia. Tal ideia se baseia no fato de que em períodos de crise, face ao comprometimento de seu fluxo de caixa, as empresas recorrem ao uso de capital de terceiros (ZEITUN; TEMIMI; MIMOUNI, 2017). Sendo assim, espera-se que, em geral, a relação entre o endividamento e o risco de insolvência tenha sido fortalecida durante o período de pandemia, o que leva às primeiras do segundo bloco de hipóteses deste estudo:

Hipótese 2a: A pandemia de COVID-19 fortalece a influência positiva da dívida líquida de curto prazo no risco de insolvência.

Hipótese 2b: A pandemia de COVID-19 fortalece a influência positiva do endividamento de curto prazo no risco de insolvência.

Hipótese 2c: A pandemia de COVID-19 fortalece a influência positiva do endividamento de longo prazo no risco de insolvência.

Conforme já mencionado, espera-se que o benefício fiscal proporcione uma relação quadrática entre o endividamento financeiro a valor contábil no risco de insolvência. Assim, espera-se que inicialmente, este endividamento influencie negativamente no risco de insolvência, mas a partir de determinado ponto, espera-se que os custos de falência façam com que a influência se torne positiva.

No que se refere à pandemia de COVID-19, espera-se que a relação inicialmente negativa entre o endividamento financeiro e a pandemia seja enfraquecido, especialmente em empresas com menor flexibilidade financeira, uma vez que o endividamento devido à necessidade de recursos para a continuidade da empresa pode aumentar o risco de rolagem da dívida (FAHLENBRACH; RAGETH; STULZ, 2021), aumentando também os custos de falência, fato que pode minimizar os efeitos do escudo fiscal, enfraquecendo a relação inicialmente negativa do endividamento financeiro no risco de insolvência. Dessa maneira, a hipótese 2d do estudo foi proposta:

Hipótese 2d: A pandemia de COVID-19 enfraquece a influência negativa (até certo ponto) do endividamento financeiro no risco de insolvência.

Apesar disso, argumenta-se que, diante da necessidade do uso de capital de terceiros durante a pandemia, face à possível necessidade do uso de capital de terceiros para manter os pagamentos em dia (ZEITUN; TEMIMI; MIMOUNI, 2017), espera-se que a pandemia fortaleça a influência positiva (a partir de certo ponto – termo quadrático) do endividamento financeiro a valor contábil no risco de insolvência. Sendo assim, uma vez que espera-se que a partir de certo ponto o endividamento financeiro a valor contábil (termo quadrático) passe a influenciar positivamente no risco de insolvência, conforme a teoria do *trade off*, espera-se que essa relação tenha sido fortalecida pela pandemia de COVID-19, conforme proposto na hipótese 2e.

Hipótese 2e: A pandemia de COVID-19 fortalece a influência positiva (a partir de certo ponto) do endividamento financeiro (termo quadrático) no risco de insolvência.

Conforme proposto pelas hipóteses levantadas, tendo como base a teoria do *trade-off*, espera-se que a relação entre as variáveis de estrutura de capital e risco de insolvência, em geral seja positiva. No entanto, considerando a teoria do *trade off*, que defende a existência do benefício fiscal proporcionado pelo chamado passivo oneroso, espera-se uma relação quadrática entre o endividamento financeiro a valor contábil e o risco de insolvência. Ademais, considerando o período de pandemia, espera-se que a COVID-19 tenha influenciado a relação entre as variáveis de estrutura de capital e risco de insolvência.

3 METODOLOGIA

Nesta seção a metodologia do estudo será detalhada. Para isso, serão apresentadas informações sobre os procedimentos de coleta de dados, amostra, variáveis e procedimentos de análise.

3.1 Procedimentos de coleta de dados

Para a análise do estudo, foram usadas informações das demonstrações financeiras de empresas brasileiras com ações negociadas na B3. Os dados foram coletados por meio do banco de dados Económica, uma plataforma que apresenta dados do mercado financeiro, fundos e ações de empresas da América Latina e dos Estados Unidos. Sendo assim, os dados das empresas brasileiras não financeiras com ações negociadas na B3 foram baixados por meio da plataforma e, posteriormente organizados por meio do M.S. Excel (2010), por meio do qual os indicadores econômico-financeiros e demais variáveis foram calculados e organizados.

O período de análise escolhido foi de janeiro de 2010 a dezembro de 2021. O ano de 2010 foi o inicial devido às alterações na legislação contábil, pelas Leis 11.638/07 e 11.941/09, que, de acordo com Freire *et al.* (2012), alteraram a Lei das Sociedades Anônimas (Lei 6.404/76) para a padronização das normas contábeis brasileiras à legislação internacional. Já o ano de 2021 foi ano de encerramento, pois ele é o último ano que possui dados anuais disponíveis.

3.2 Amostra

A amostra do estudo foi composta por empresas brasileiras não financeiras de capital aberto com ações negociadas na B3. Para a coleta na base de dados Económica, foram filtradas as empresas do Brasil, que negociam ações na B3 e encontram-se ativas. Na filtragem inicial, foram excluídas também empresas classificadas do setor de finanças, de acordo com o setor econômica, ou do setor financeiro, de acordo com o setor econômico Bovespa.

Inicialmente, o número de empresas que se adequaram à filtragem foi de 563 empresas. No entanto, devido à existência de dados faltantes, o total de empresas analisadas foi de 510. Ressalta-se também que um painel desbalanceado de dados foi utilizado na análise, totalizando 15.582 observações. Ademais, ressalta-se que foram analisados dados trimestrais das empresas, totalizando 48 trimestres no período de 2010 a 2021.

3.3 Variáveis

Para a elaboração do estudo, os indicadores econômico-financeiros usados foram obtidos por meio do Economática. Para o tratamento das variáveis, utilizou-se a *winsorização* de até 5% (2,5% em cada cauda: superior e inferior), para a eliminação de *outliers*, que podem prejudicar a análise (HAIR, *et al.*, 2006), e alcançar valores de assimetria e curtose próximos a zero e três, respectivamente (FÁVERO; BELFIORE, 2017).

Ademais, para evitar a ocorrência de multicolinearidade entre as variáveis do modelo utilizou-se o fator de inflação da variância (VIF) e a matriz de correlação. De acordo com Gujarati e Porter (2011), é normal que as variáveis explicativas possuam correlação, mas o problema de multicolinearidade existe quando essa correlação é alta. Segundo Gujarati e Porter (2011), valores do VIF superior a 10 e de correlações maiores que 0,8 entre as variáveis indicam a existência de multicolinearidade.

A variável dependente escolhida foi o *Z-score* do trabalho de Altman, Baidya e Dias (1979), no qual foi desenvolvido um modelo de classificação e previsão de problemas financeiros para empresas brasileiras. No estudo, os autores se basearam no modelo de Altman (1968), mas adaptaram-no para as condições brasileiras. Sendo assim, o risco de insolvência será calculado considerando os parâmetros da Equação 4:

$$Z_1 = -1,44 + 4,03X_2 + 2,25X_3 + 0,14X_4 + 0,42X_5 \quad (4)$$

Em que:

$$X_2 = \frac{\text{Não exigível} - \text{Capital Aportado pelos Acionistas}}{\text{Ativo Total}} \quad (5)$$

$$X_3 = \frac{\text{Lucro Antes dos Juros e Impostos}}{\text{Ativo Total}} \quad (6)$$

$$X_4 = \frac{\text{Valor Contábil do Patrimônio Líquido}}{\text{Exigível Total}} \quad (7)$$

$$X_5 = \frac{\text{Vendas}}{\text{Ativo Total}} \quad (8)$$

De acordo com os autores, o ponto crítico de separação das empresas com e sem problemas (valor de Z_1) é zero. Isso significa que as empresas com Z_1 menor que zero podem ser consideradas empresas que apresentam problemas sérios, ou seja, maior probabilidade de falência. Já as empresas com Z_1 maior que zero são empresas que têm menor probabilidade de falência. Nesse caso, da forma como foi proposto, quanto maior o índice Z_1 menor a probabilidade de falência. Para fins de um melhor entendimento e para que quanto maior o valor de Z_1 , maior a probabilidade de insolvência, o valor de Z_1 foi multiplicado por menos um (-1). Assim, com essa transformação, quanto maior o Z_1 maior será a probabilidade de falência das empresas da análise.

Para a análise proposta no estudo, foram consideradas variáveis de endividamento como explicativas, para analisar os efeitos da estrutura de capital no risco de insolvência. As variáveis de endividamento usadas e os sinais esperados dos coeficientes foram detalhados na Tabela 8.

Tabela 8 — Variáveis independentes do estudo.

Autores	Indicador	Fórmula	Sigla	Sinal esperado
Fahlenbrach, Rageth e Stulz (2021)	Dívida líquida de curto prazo	Dívida líquida de curto prazo total = Passivo Circulante – (Caixa e equivalente de Caixa + Aplicações financeiras)	<i>DIV_LIQ_CP</i>	+
Baker e Wurgler (2002) e Albanez (2012), Prado (2019).	Endividamento financeiro	$\frac{\text{Dívida financeira bruta total}}{\text{Ativo total}}$	<i>END_C_DF</i>	-
		Dívida financeira bruta total = passivo oneroso	<i>END_C_DF_2</i>	+
Forte (2007) e Bastos e Nakamura (2009), Prado (2019).	Endividamento de curto prazo	$\frac{\text{Passivo circulante}}{\text{Ativo total}}$	<i>END_C_CP</i>	+
Forte (2007) e Bastos e Nakamura (2009), Prado (2019).	Endividamento de longo prazo	$\frac{\text{Passivo não circulante}}{\text{Ativo total}}$	<i>END_C_LP</i>	+

Fonte: Da autora (2023).

A primeira variável do estudo foi a dívida líquida de curto prazo (DIV_LIQ_CP). O cálculo dessa variável foi realizado por meio da subtração entre o passivo circulante e o valor em caixa e em aplicações financeiras da empresa, que mostra o percentual da dívida de curto prazo que excede o caixa da empresa, representado pelas contas de alta liquidez do ativo (FAHLENBRACH; RAGETH; STULZ, 2021). Essa variável foi usada para a análise da primeira hipótese do primeiro bloco (1a), que propôs uma influência positiva no risco de insolvência, de acordo com a teoria do *trade-off*.

A variável de endividamento de curto prazo a valor contábil foi umas das variáveis de estrutura de capital considerada no modelo. Seu cálculo foi realizado por meio da divisão entre as dívidas de curto prazo (passivo circulante) e o ativo total (BASTOS; NAKAMURA; BASSO, 2009; PRADO, 2019). Essa variável foi usada para a análise da segunda hipótese do primeiro bloco (1b), que de acordo com a teoria do *trade-off*, propôs uma influência positiva no risco de falência.

O endividamento de longo prazo também foi considerado no modelo. Essa variável indica o nível de endividamento considerando fontes de recursos de longo prazo. Seu cálculo pode ser realizado por meio da razão entre o passivo não circulante e o ativo total (BASTOS; NAKAMURA, 2009; PRADO, 2019). Essa variável foi usada para testar a hipótese 1c, que é consistente com a teoria do *trade-off*, propôs um efeito negativo no risco de insolvência.

Outra variável do estudo foi o endividamento financeiro a valor contábil. Essa variável foi calculada por meio da razão entre o passivo oneroso (total de empréstimos e financiamentos de curto e longo prazo) e ativo total (ALBANEZ; VALLE; CORRAR, 2012; BAKER; WURGLER, 2002; PRADO, 2019). Coerente com a teoria do *trade-off*, que postula sobre a existência do benefício fiscal, espera-se que as despesas financeiras provenientes do passivo oneroso inicialmente influenciam negativamente no risco de falência.

No entanto, devido à existência dos custos de falência, espera-se que essa influência seja apenas até certo ponto, existindo, portanto uma relação quadrática em forma de U, conforme proposto na hipótese 1d. Assim, a variável foi elevada ao quadrado para verificar a relação não monotônica. De acordo com a hipótese 1d, espera-se que o endividamento financeiro a valor contábil apresente um coeficiente significativo e com sinal negativo e que o termo quadrático dessa variável apresente um coeficiente significativo e com sinal positivo.

Uma vez que a literatura sugere que fatores do ambiente externo da organização influenciam na estrutura de capital e no risco de falência (CARVALHO; CURTO; PRIMOR, 2022; HACKBARTH; MIAO; MORELLEC, 2006; TINOCO; WILSON, 2013), uma variável *dummy* para a COVID-19 também foi considerada no modelo. Uma vez que de acordo com o

Ministério da Saúde (2020), o primeiro caso registrado de COVID-19 foi em 26 de fevereiro de 2020, a variável *dummy* recebeu o valor um (1) para os trimestres referentes aos anos de 2020 e 2021 e zero (0) para os demais trimestres da amostra.

Para analisar o efeito da pandemia na relação entre estrutura de capital e risco de insolvência, foram criadas variáveis com interação por meio da multiplicação entre a variável *dummy* e as variáveis de estrutura de capital. Diante disso, caso as variáveis com interação possuam resultados significativos, interpreta-se que houve uma influência da pandemia de COVID-19 na relação entre estrutura de capital e risco de falência, em que coeficientes positivos indicam que houve um fortalecimento da relação e coeficientes negativos indicam o enfraquecimento da relação entre estrutura de capital e risco de falência.

3.4 Procedimentos de análise

Para a análise proposta no estudo foi utilizada a técnica de regressão multinível de dados em painel. Os dados em painel são entendidos como um conjunto de dados contendo observações de séries temporais e de vários indivíduos ao mesmo tempo (HSIAO, 2007).

Gujarati e Porter (2011) afirmam que os dados em painel são uma combinação entre dados transversais (que incluem uma única observação de diversos indivíduos em um único período) e séries temporais (que incluem diversas observações de um único indivíduo em diferentes períodos). Os dados transversais podem ser entendidos como dados em que se tem uma única observação de vários indivíduos em apenas um período de tempo (mês, trimestre, ano, entre outros). Já as séries temporais apresentam diversas observações de apenas um indivíduo em diferentes períodos.

Entre as vantagens do uso da regressão de dados em painel, Gujarati e Porter (2011) mencionam a maior variabilidade, existência de mais graus de liberdade, mais eficiência e menos colinearidade entre variáveis. Hsiao (2007) complementa afirmando que os dados em painel possuem vantagens se comparados aos dados transversais e séries temporais. Entre elas estão a maior precisão na inferência dos parâmetros do modelo, capacidade maior para capturar a complexidade dos indivíduos e, em alguns casos, o cálculo e a inferência estatística são simplificados.

Existem, no entanto, diferentes tipos de modelos para realização de análises de dados em painel. Neste estudo, optou-se pelo uso de um modelo multinível de regressão para dados em painel. De acordo com Fávero e Belfiore (2017), os modelos multinível de regressão para dados em painel tem sido cada vez mais utilizados em diferentes áreas do conhecimento.

De acordo com os mesmos autores, nesses modelos são consideradas estruturas aninhadas de dados, ou seja, que existem grupos entre as observações que apresentam características em comum, como setores, por exemplo. Sendo assim, é possível que o aninhamento natural existente nos dados seja levado em consideração, o que permite que heterogeneidades individuais e entre grupos sejam analisadas, e portanto, seja possível especificar os componentes aleatórios dos diferentes níveis de análise.

A escolha do modelo multinível se deu face às análises especificadas por Fávero e Belfiore (2017). Sendo assim, foram desenvolvidos um modelo nulo, um modelo com interceptos aleatórios e um modelo com interceptos e inclinações aleatórios. Após os resultados dessas análises, optou-se por um modelo hierárquico linear de três níveis com interceptos e inclinações aleatórias. Foram determinados três níveis pois os dados variam no tempo (nível 1), entre empresas (nível 2) e entre setores (nível 3). Assim, os períodos de tempo encontram-se aninhados (ou agrupados) em empresas e as empresas encontram-se aninhadas (ou agrupadas) em setores.

O modelo nulo, de acordo com Fávero e Belfiore (2017), permite para o presente estudo, a verificação da existência de variabilidade do risco entre empresas do mesmo setor e entre empresas de setores distintos. Nesse modelo, nenhuma variável explicativa é inserida, ou seja, considera um intercepto e os termos de erro. Este modelo pode ser representado pela Equação 10.

$$Risco_{tjk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{tjk} \quad (10)$$

Em que:

$Risco_{tjk}$ = Variável dependente (Risco de insolvência)

γ_{000} = intercepto geral (constante que não varia entre os indivíduos);

u_{00k} = efeitos aleatórios do nível 3 - setor;

r_{0jk} = efeitos aleatórios do nível 2 - empresa

e_{tjk} = termos de erro de nível 1 – tempo (trimestre).

Após a definição do modelo nulo, estimou-se as correlações intraclasse de nível 2 e nível 3. A correlação intraclasse mede a variância total proveniente dos níveis 2 e 3 (FÁVERO; BELFIORE, 2017). Para o cálculo, utiliza-se, respectivamente:

$$rho_{empresa/sector} = Corr(Y_{tjk}, Y_{t'jk}) = \frac{\tau_{u000} + \tau_{r000}}{\tau_{u000} + \tau_{r000} + \sigma^2} \quad (11)$$

$$rho_{sector} = Corr(Y_{tjk}, Y_{t'jk}) = \frac{\tau_{u000}}{\tau_{u000} + \tau_{r000} + \sigma^2} \quad (12)$$

Em que:

$rho_{empresa/sector}$ = correlação intraclasse de nível 2;

rho_{sector} = correlação intraclasse de nível 3;

τ_{u000} e τ_{r000} = variâncias nos termos de erro;

σ^2 = variância residual nos termos de erro.

Se a correlação intraclasse for igual a zero, indica que não existe variância entre os grupos (nível 2 ou 3). Se for diferente de zero, indica que existe variância entre os grupos, logo o modelo multinível se torna mais adequado que modelos tradicionais, como Mínimos Quadrados Ordinários (MQO), por exemplo (FÁVERO; BELFIORE, 2017). Sendo assim, a correlação intraclasse de nível 2 se refere à correlação entre os riscos de insolvência trimestrais, para uma mesma empresa de determinado setor, enquanto a de nível 3 se refere à correlação entre os riscos de insolvência trimestrais, para um mesmo setor.

Ademais, verificou-se o resultado do teste de razão de verossimilhança, que também indica se um modelo multinível é o mais adequado. Como resultado, percebeu-se que o p-valor foi significativo ($0,000 < 0,05$), o que indica que o multinível seja mais adequado que um modelo tradicional.

Após a identificação de que um multinível seria o mais adequado para as análises do presente estudo, foram estimados um modelo com interceptos aleatórios e um modelo com interceptos e inclinações aleatórias, para decidir por meio dos valores das correlações intraclasse e do resultado do teste de razão de verossimilhança. Após a elaboração dos modelos, verificou-se que o mais adequado seria o modelo com interceptos e inclinações aleatórias. Assim, o modelo estimado no presente estudo está representado na Equação 13:

$$\begin{aligned}
Risco_{tjk} = & \gamma_{000} + \gamma_{100} \cdot Trimestre_{jk} + \gamma_{010} \cdot DIV_LIQ_CP_{jk} \\
& + \gamma_{010} \cdot END_C_CP_{jk} + \gamma_{010} \cdot END_C_LP_{jk} \\
& + \gamma_{010} \cdot END_C_DF_{jk} + \gamma_{010} \cdot END_C_DF_2_{jk} \gamma_{010} \cdot KI_{jk} \\
& + \gamma_{110} \cdot DIV_LIQ_CP_{jk} \cdot Trimestre_{jk} \\
& + \gamma_{110} \cdot END_C_CP_{jk} \cdot Trimestre_{jk} \\
& + \gamma_{110} \cdot END_C_LP_{jk} \cdot Trimestre_{jk} \\
& + \gamma_{110} \cdot END_C_DF_{jk} \cdot Trimestre_{jk} \\
& + \gamma_{110} \cdot END_C_DF_2_{jk} \cdot Trimestre_{jk} + u_{00k} \\
& + u_{10k} \cdot Trimestre_{jk} + r_{0jk} + r_{1jk} \cdot Trimestre_{jk} + e_{tjk}
\end{aligned} \tag{13}$$

Para a realização da estimação do modelo utilizou-se o *software* Stata, versão 14.2. Os resultados obtidos por meio da análise foram detalhados na seção de resultados, a seguir.

4 RESULTADOS

Antes de iniciar as análises da regressão multinível dos dados em painel, os dados foram tratados por meio da winsorização. Esse processo tem como objetivo a remoção de *outliers*, que prejudicam as análises. Sendo assim, as variáveis do estudo foram winsorizadas considerando o máximo de 5% das observações em cada variável (na maioria dos casos seguindo o critério de 2,5% para cada cauda), tentando também manter valores de assimetria e curtose próximos de zero (0) e três (3), respectivamente (FÁVERO; BELFIORE, 2017).

A Tabela 9 mostra a estatística descritiva das variáveis do estudo, assim como os valores de assimetria e curtose de cada variável após o tratamento dos dados.

Tabela 9 — Estatística descritiva.

Variáv.	Méd.	Máx.	Mín.	D. Padrão	Var.	Ass.	Curt.
DIV_LIQ_CP	1015694	7700000	-4100000	2,024,812	4.10e+12	1.997	7.142
END_C_CP	0.295	1.629	0.005	0.301	0.090	2.830	12.136
END_C_LP	0.427	2.173	0	0.390	0.152	2.765	12.510
END_C_DF	0.337	1.189	0	0.258	0.066	1.066069	4.637

Legenda: DIV_LIQ_CP = dívida líquida de curto prazo; END_C_CP = endividamento de curto prazo; END_C_LP = endividamento de longo prazo; END_C_DF = endividamento financeiro.

Fonte: Da autora (2023).

Conforme percebe-se, a variável com maior valor de assimetria foi a variável END_C_CP, com um valor igual a 2.830472. A variável com maior valor de curtose, por sua vez, foi a variável END_C_LP, com um valor de curtose igual a 12.51064. No que se refere à média, aquela com o maior valor médio também foi a DIV_LIQ_CP, de 1,015,694 no período. Em seguida estiveram as variáveis END_C_LP, END_C_DF e END_C_CP, com valores médios de 0.4265618, 0.3365776 e 0.2954854, respectivamente.

Ademais, para verificar a existência de multicolinearidade entre as variáveis explicativas analisou-se, primeiramente, o fator de inflação da variância (VIF). De acordo com Gujarati e Porter (2011), caso o VIF apresente valores superiores a 10, existe o problema de multicolinearidade entre as variáveis explicativas do modelo. Sendo assim, as variáveis devem apresentar valores de fator de inflação da variância menores que 10. A Tabela 10 mostra os valores do fator de inflação da variância das variáveis explicativas do usadas no estudo.

Tabela 10 — Valores do fator de inflação da variância (continua).

Variável	VIF
DIV_LIQ_CP	1.04
END_C_CP	1.09

Tabela 10 — Valores do fator de inflação da variância (conclusão).

Variável	VIF
END_C_LP	1.36
END_C_DF	1.41

Fonte: Da autora (2023).

Conforme percebeu-se por meio da Tabela 10, o valor máximo de fator de inflação da variância foi de 1.41, referente à variável END_C_DF. Isso indica que todas as variáveis do modelo apresentaram valores menores do que o recomendado por Gujarati e Porter (2011), ou seja, valores menores que 10. Logo, esses valores indicam a não existência do problema de multicolinearidade entre as variáveis do estudo.

Para confirmar tal constatação, analisou-se também a matriz de correlação. Por meio dela, também é possível identificar o problema de multicolinearidade entre as variáveis do estudo. Gujarati e Porter (2011) afirmam que correlações maiores que 0,8 indicam multicolinearidade. A Tabela 11 mostra os valores das correlações entre as variáveis.

Tabela 11 — Matriz de correlação.

Variável	DIV_LIQ_CP	END_C_CP	END_C_LP	END_C_DF
DIV_LIQ_CP	1			
END_C_CP	0.1866	1		
END_C_LP	-0.0551	0.1195	1	
END_C_DF	0.0392	0.2302	0.5104	1

Fonte: Da autora (2023).

Conforme observado na Tabela 11, o maior valor de correlação foi de 0.5104, entre as variáveis END_C_DF e END_C_LP. Esse valor se encontra abaixo de 0,8, valor considerado como indício da existência do problema de multicolinearidade. Dessa maneira, a matriz de correlação reforça a não existência de alta correlação entre as variáveis explicativas do estudo.

Em seguida, verificou-se a melhor adequação do modelo multinível em relação a modelos tradicionais de regressão de dados em painel. Para isso, os passos sugeridos por Fávero e Belfiore (2017) foram seguidos. Primeiramente, estimou-se o modelo nulo, que de acordo com os autores mostra se o modelo multinível é mais adequado que modelos tradicionais, principalmente por meio da correlação intraclasse. A correlação intraclasse mostra quanto da variância total do modelo se deve aos diferentes níveis do modelo. Sendo assim, valores de correlação intraclasse diferentes de zero indicam que o modelo multinível é mais adequado.

A partir da estimação do modelo nulo, chegou-se aos resultados mostrados na Tabela 12.

Tabela 12 — Resultados do modelo nulo.

Efeitos fixos			
Constante			
Item	Coefficiente Beta	Erro Padrão	P-valor
_cons	2.036976	0.3024864	0.000
Efeitos aleatórios			
Item	Coefficiente	Erro Padrão	τ_{xxxx} / Erro-padrão
τ_{u000}	0.4684731	0.3731172	1.2555655435
τ_{r000}	24.22173	1.544636	15.6811896136
σ^2	7.407729	0.0831774	89.0593959416
Correlação Intraclasse			
Item	ICC	Erro Padrão	
rhoempresa setor	0.7692148	0.0114856	
rhoetor	0.0145951	0.0115063	
Teste de Razão de Verossimilhança			
	Chi2 (2)	P-valor	
LR-teste	25495.18	0.000	

Fonte: Adaptado de Pereira (2021).

Por meio dos resultados do modelo nulo (Tabela 12), percebe-se que o componente de efeitos fixos γ_{000} é de 2,0369. Este parâmetro indica o intercepto geral do estudo, ou seja, a média dos riscos de insolvência trimestrais esperados das empresas. Ademais, as variâncias dos termos de erro τ_{u000} , τ_{r000} e σ^2 foram de 0,4684731, 24,22173 e 7,407729, respectivamente (no Stata correspondem aos parâmetros $\text{var}(_cons)$ para Setor, $\text{var}(_cons)$ para empresa e $\text{var}(\text{Residual})$).

A partir desses valores, as correlações intraclasse de nível 2 e nível 3 foram de 76,92% e 1,45%, respectivamente. Isso indica que a correlação entre os riscos de insolvência trimestrais para um mesmo setor é de aproximadamente 1,45%. Já a correlação entre os riscos de insolvência trimestrais para uma mesma empresa de determinado setor é de 76,92%. Assim, considerando o modelo nulo (sem variáveis explicativas), o risco de insolvência trimestral tem baixa correlação entre setores, mas o risco de insolvência trimestral tem correlação considerável para a mesma empresa de um determinado setor. Logo, estima-se que os efeitos aleatórios de empresas e setores correspondem a cerca de 76,92% do total de variância dos resíduos.

A significância estatística das variâncias foi confirmada pela divisão das variâncias dos termos de erro τ_{u000} , τ_{r000} e σ^2 por seus respectivos erros-padrão. Essa razão, se maior que 1,96, corresponde à um nível de significância de 5%. Para esses casos, esses valores foram de 1,25%, 15,68 e 89,15, o que indica que um dos valores foi menor que 1,96.

Para confirmar a existência de variação significativa no risco de insolvência anual entre empresas e setores, deve-se observar o resultado do teste de razão de verossimilhança, que foi mostrado na parte inferior da Tabela 12. O resultado do teste a rejeitou a hipótese nula de que os interceptos aleatórios são iguais a zero (p -valor=0,000). Sendo assim, confirmou-se que um modelo multinível é mais indicado para este estudo, se comparado a modelos tradicionais.

Uma vez determinado que um modelo de regressão multinível é mais adequado, é preciso analisar se deve-se escolher um modelo de tendência linear com interceptos aleatórios ou um modelo de tendência linear com interceptos e inclinações aleatórias. Para isso, faz-se necessário estimar ambos os modelos, inserindo no primeiro caso a variável de nível 1 (trimestre) e, no segundo caso, inserir os efeitos aleatórios de inclinação nos níveis 2 e 3.

Assim, foi possível identificar a existência de fatores das empresas e dos setores responsáveis pela variação do risco de insolvência entre empresas do mesmo setor e de setores distintos, bem como identificar a existência de variância nas inclinações do risco de insolvência no decorrer do tempo, entre as empresas diferentes.

Primeiramente, estimou-se o modelo de tendência linear com interceptos aleatórios, que apresentou os resultados mostrados na Tabela 13:

Tabela 13 — Modelo de tendência linear com interceptos aleatórios.

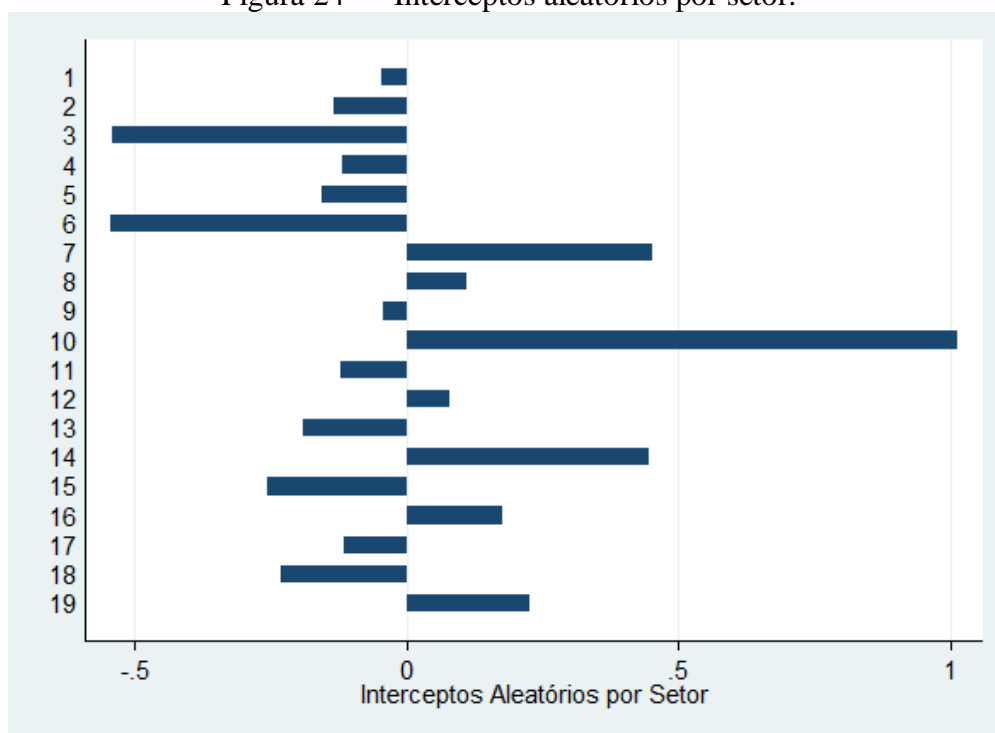
Efeitos fixos				
Medida Repetida no nível 1				
Descrição	Variável	Coefficiente Beta	Erro Padrão	P-valor
	Trimestre	0.0368475	0.0016485	0.000
Constante				
Item		Coefficiente Beta	Erro Padrão	P-valor
_cons		0.8667207	0.3117598	0.005
Efeitos aleatórios				
Item		Coefficiente	Erro Padrão	$\tau_{xxx}/$ Erro-padrão
τ_{u000}		0.4911108	0.3976983	1.2348828245
τ_{r000}		24.78676	1.581865	15.6693270285
σ^2		7.176917	0.080592	89.052474191
Correlação Intraclasse				
Item		ICC		Erro Padrão
$\rho_{\text{empresa setor}}$		0.7788641		0.0111556
ρ_{setor}		0.0151322		0.0121261
Teste de Razão de Verossimilhança				
		Chi2 (2)		P-valor
LR-teste		25963.12		0.000

Fonte: Adaptado de Pereira (2021).

Por meio dos resultados, percebe-se que a média de crescimento trimestral do risco de insolvência foi estatisticamente significativo, com valor de $\gamma_{100} = 0,0368475$. Novamente, por meio dos componentes de efeitos aleatórios (termos de erro τ_{u000} , τ_{r000} e σ^2), foi possível estimar as correlações intraclasse de nível 2 e nível 3. Os valores foram de 77,88% e 1,51%, respectivamente. Os valores das correlações intraclasse de nível 2 e nível 3 nesse caso foram superiores às do modelo nulo (76,92% e 1,45%). Isso indica que é importante inserir a variável de medida repetida no nível 1 (que no caso deste estudo, é o trimestre).

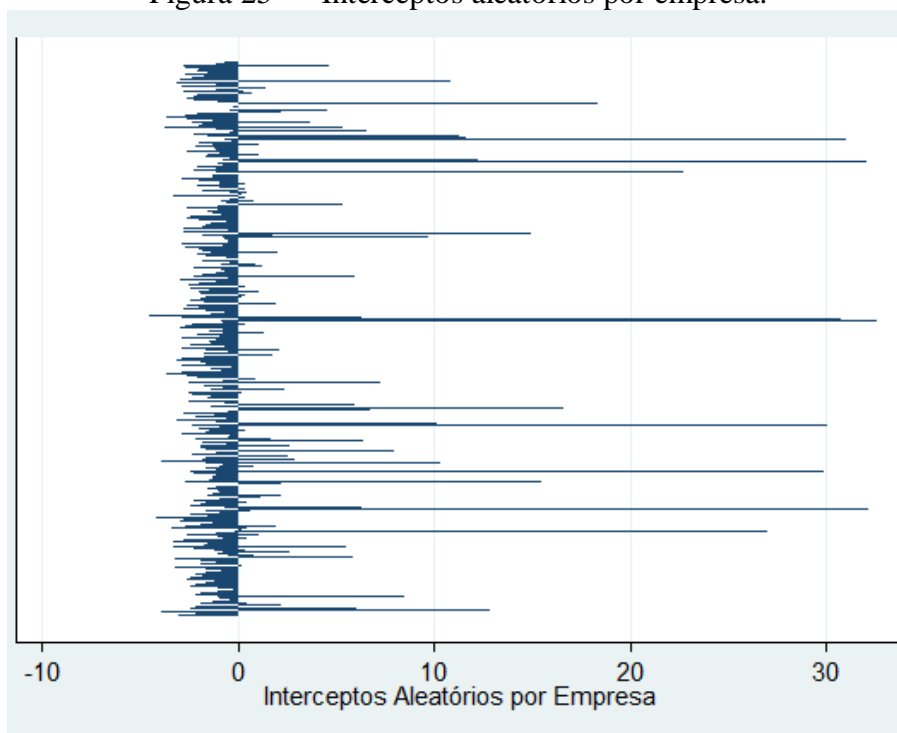
Para melhor visualizar os interceptos aleatórios por setor e por empresa, pode-se observar as Figuras 24 e 25:

Figura 24 — Interceptos aleatórios por setor.



Fonte: Da autora (2023).

Figura 25 — Interceptos aleatórios por empresa.



Fonte: Da autora (2023).

Por meio das Figuras 24 e 25 é possível observar que embora os termos de erro u_{00k} , não variam para empresas do mesmo setor ao longo do tempo, os termos r_{0jk} variam entre empresas, mas não para uma mesma empresa ao longo do tempo. Tal situação caracteriza um intercepto para cada empresa e um intercepto para cada setor. Esses resultados reforçam, mais uma vez, o uso de um modelo multinível.

Por fim, estimou-se um modelo de tendência linear com interceptos e inclinações aleatórias, para identificar se existe também variância significativa nas inclinações do risco de insolvência ao longo do tempo entre as empresas. Os resultados do modelo estão detalhados na Tabela 14.

Tabela 14 — modelo de tendência linear com interceptos e inclinações aleatórias (continua).

Efeitos fixos				
Medida Repetida no nível 1				
Descrição	Variável	Coefficiente Beta	Erro Padrão	P-valor
	Trimestre	0.0285313	0.0078202	0.000
Constante				
Item		Coefficiente Beta	Erro Padrão	P-valor
	_cons	1.357991	.4269353	0.001

Tabela 14 — modelo de tendência linear com interceptos e inclinações aleatórias (conclusão).

Efeitos aleatórios			
Item	Coefficiente	Erro Padrão	τ_{xxxx} / Erro-padrão
τ_{u000}	0.7471923	0.9636276	0.7753952876
τ_{u100}	5.74e-12	6.42e-11	0.0894081
τ_{r000}	45.75911	3.413035	13.4071610751
τ_{r100}	0.025599	0.0018759	13.6462498001
σ^2			
Correlação Intraclasse			
Item	ICC	Erro Padrão	
$\rho_{\text{empresa setor}}$	0.9172623	0.005721	
ρ_{setor}	0.0147372	0.0188627	
Teste de Razão de Verossimilhança			
	Chi2 (2)	P-valor	
LR-teste	32730.20	0.000	

Fonte: Adaptado de Pereira (2021).

Por meio dos resultados, as correlações intraclasse de nível 2 e 3 foram de 91,72% e 1,47%. Assim, espera-se que 91,72% da variância total dos resíduos seja explicada pelos efeitos aleatórios de empresas e setores. Como pode-se observar, a correlação intraclasse de nível 2 superou a do modelo de tendência linear com interceptos aleatórios (77,88%), mas não superou a correlação intraclasse de nível 3 (1,51%). Para confirmar qual o modelo mais adequado (modelo de tendência linear com interceptos ou modelo de tendência linear com interceptos e inclinações aleatórias), realizou-se o teste de razão de verossimilhança (lrtest). Por meio do teste, é possível verificar a necessidade de inclusão de inclinações aleatórias, ou seja, se o modelo deve ser com interceptos e inclinações aleatórias.

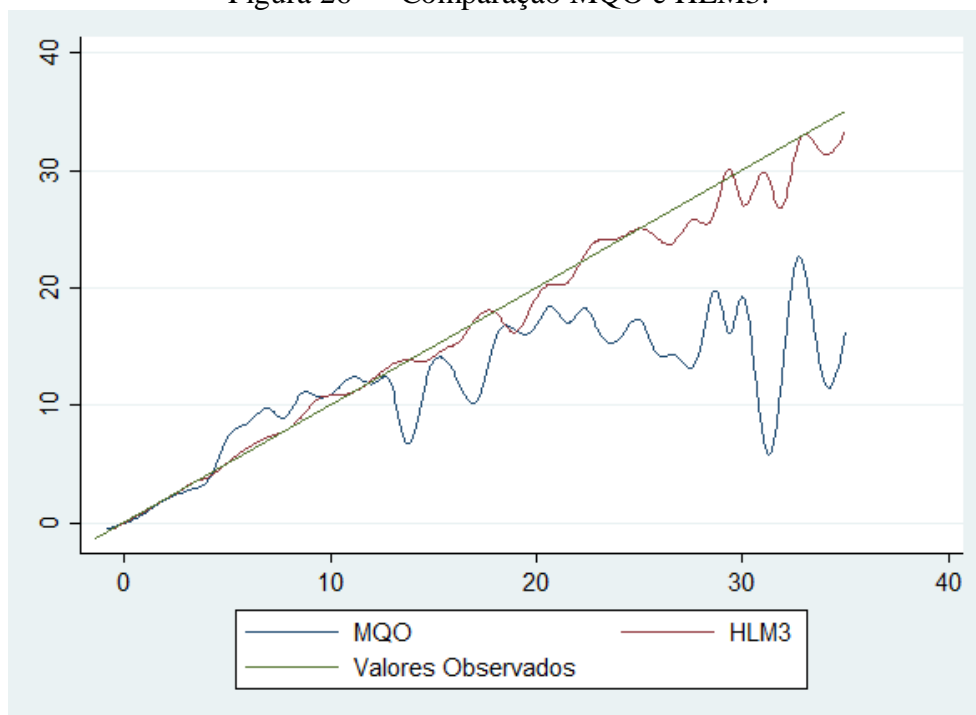
O resultado do teste de razão de verossimilhança (lrtest) foi significativo (p-valor = 0,000), mostrando que um modelo de tendência linear com interceptos e inclinações aleatórias é o mais adequado, o que implica que deve-se considerar um modelo com interceptos e inclinações aleatórias no presente estudo.

Após isso, as variáveis explicativas do estudo foram incluídas, no entanto, foi necessário analisar a estrutura das matrizes de variância-covariância dos efeitos aleatórios (FÁVERO; BELFIORE, 2017). Inicialmente nenhuma estrutura de covariância dos termos de erro foi especificada. Assim, no modelo inicial espera-se que $\text{cov}(u_{00k}, u_{10k}) = 0$ e $\text{cov}(r_{0jk}, r_{1jk}) = 0$. Mas é possível generalizar a estrutura das matrizes de covariância permitindo que u_{00k} e u_{10k} sejam correlacionados, assim como r_{0jk} e r_{1jk} .

Para isso, adicionou-se ao modelo o termo *covariance (unstructured)* nos componentes de efeitos aleatórios do nível setor e no nível empresa. Após isso, para decidir se o modelo com o termo *covariance (unstructured)* é mais adequado, realizou-se um teste de razão de verossimilhança, com o intuito de comparar a estimação inicial com a estimação considerando a estrutura de covariância. O resultado do teste ($Qui\ quadrado=0,000<0,05$) indicou que a estrutura das matrizes de covariância variância-covariância dos efeitos aleatórios pode ser considerada *unstructured*, isto é, os termos de erro podem ser considerados correlacionados.

Por fim, para mostrar visualmente a comparação de um modelo tradicional (MQO – Mínimos Quadrados Ordinários) com o modelo multinível (HLM3) para esse caso, elaborou-se o gráfico mostrado na Figura 26.

Figura 26 — Comparação MQO e HLM3.



Fonte: Da autora (2023).

Como pode ser observado na Figura 26, o modelo multinível (HLM3) apresentou-se mais próximo aos valores observados. Isso mostra, visualmente, a superioridade do modelo de tendência linear com variáveis explicativas e com interceptos e inclinações aleatórias nos níveis 2 e 3 (HLM3) se comparado ao modelo de regressão linear múltipla estimado por MQO. Assim, menciona-se a relevância de considerar os componentes de efeitos aleatórios em estruturas de dados aninhadas.

A partir da definição do modelo multinível mais adequado para a análise (modelo de tendência linear com interceptos e inclinações aleatórias), as variáveis explicativas do estudo foram adicionadas ao modelo, para a análise das hipóteses do estudo. Conforme já mencionado, as variáveis explicativas do estudo foram variáveis que indicam a estrutura de capital das empresas, com o intuito de verificar o efeito no risco de insolvência. Na Tabela 15 encontram-se os resultados do modelo completo.

Tabela 15 — Resultados do modelo completo (continua).

Efeitos fixos					
Medida Repetida no nível 1					
Descrição	Variável	Coeficiente Beta	Erro Padrão	P-valor	
	Trimestre	-0.0058043	0 .0097534	0.552	
Hipóteses					
Hipótese	Variável	Coeficiente Beta	Erro Padrão	P-valor	Conclusão da Hipótese
1a	DIV_LIQ_C P	-2.12e-07	3.77e-08	0.000	Não Confirma
1b	END_C_CP	5.84824	0 .2437274	0.000	Confirma
1c	END_C_LP	4.936831	0 .219609	0.000	Confirma
1d	END_C_DF	-3.756267	0 .8991238	0.000	Confirma
	END_C_DF_ 2	4.013881	0 .854237	0.000	
-	DIV_LIQ_C P_ANO	2.48e-09	1.17e-09	0.033	-
-	END_C_CP_ TRIM	0 .0432201	.0079983	0.000	-
-	END_C_LP_ TRIM	0.0510475	0 .0069765	0.000	-
-	END_C_DF_ TRIM	-0.0498824	0.0269279	0.064	-
-	END_C_DF_ 2_TRIM	-0.0206045	0.0258778	0.426	-
Constante					
Item		Coeficiente Beta	Erro Padrão	P-valor	
_cons		-1.226888	0 .4080072	0.003	
Efeitos aleatórios					
Item		Coeficiente	Erro Padrão	τ_{xxx} / Erro-padrão	
τ_{000} var(_cons) para Setor		0.1586176	0.3685847	0.4303423	
τ_{100} (var(data) para setor		0.0000174	0.000081	0.2148148	
τ_{r000} var(_cons) para empresa		45.11214	3.726063	12.107187	
τ_{r100} (var(data) para empresa		0.0237997	0.0019272	12.349367	
σ^2 var(Residual))		2.778901	0.0326163	8.52E-06	

Tabela 16 — Resultados do modelo completo (conclusão).

Correlação Intraclasse		
Item	ICC	Erro Padrão
rhoempresa setor	0.9421661	0.0045477
rhoetor	0.0033011	0.0076707
Teste de Razão de Verossimilhança		
	Chi2 (2)	P-valor
LR-teste	27695.06	0.000

Fonte: Adaptado de Pereira (2021).

A partir dos resultados, ao adicionar as variáveis explicativas, o modelo completo apresentou correlações intraclasse de nível 2 e 3 iguais a 94,21% e 0,33 %. Ademais, o resultado do teste de razão de verossimilhança (lrtest) mostrou, mais uma vez a adequação do modelo multinível para o presente estudo. Assim, ao analisar a correlação intraclasse, apesar de a correlação entre o risco de insolvência para um mesmo setor ser baixa (0,33%), a correlação entre o risco de insolvência para uma mesma empresa de um mesmo setor apresentou um valor elevado (94,21%).

Para a análise da hipótese 1a do estudo (a dívida líquida de curto prazo influencia positivamente o risco de insolvência), os resultados da variável dívida líquida de curto prazo (DIV_LIQ_CP) foram observados. Por meio dos resultados, percebeu-se que a variável apresentou um resultado significativo (p-valor=0,001), porém negativo, diferente do esperado. Logo a hipótese 1a do estudo não foi confirmada. Isso indica que a dívida líquida de curto prazo influencia negativamente no risco de insolvência de empresas brasileiras não financeiras com ações negociadas na B3.

Este resultado pode ser considerado consistente com a ideia defendida por Miller (1974). De acordo com o autor, os custos de falência existem, mas podem ser proporcionalmente irrelevantes se comparados ao benefício fiscal do endividamento. Sendo assim, o benefício fiscal resultante do passivo oneroso de curto prazo, que está incluído no passivo circulante, pode ser o responsável por compensar os custos de falência associados à maior dívida líquida de curto prazo, diminuindo o risco de falência.

A análise da hipótese 1b do estudo (o endividamento de curto prazo influencia positivamente o risco de insolvência), se baseou nos resultados da variável endividamento de curto prazo (END_C_CP). Conforme apresentado na Tabela 15, o resultado dessa variável foi significativo (p-valor=0,000) e positivo. Tal resultado confirmou a hipótese 1b, indicando a influência positiva do endividamento de curto prazo no risco de insolvência das empresas

brasileiras não financeiras com ações negociadas na B3, resultado consistente com a teoria do *trade-off*.

Para analisar a hipótese 1c (o endividamento de longo prazo influencia positivamente o risco de insolvência), os resultados da variável endividamento de longo prazo (END_C_LP). Os resultados da Tabela 15 mostraram que a variável apresentou um resultado significativo (p-valor=0,000) e positivo, de acordo com o esperado. Assim, o resultado confirmou a hipótese 1c do estudo, também indicando a influência positiva do endividamento de curto prazo no risco de insolvência das empresas brasileiras não financeiras com ações negociadas na B3.

Conforme proposto pela teoria do *trade-off*, esperava-se que o nível de endividamento influenciasse positivamente no risco de falência, devido aos aumentos dos custos de falência mediante maior endividamento (MYERS, 2001). As variáveis dívida líquida de curto prazo (DIV_LIQ_CP), endividamento de curto prazo (END_C_CP) e endividamento de longo prazo (END_C_LP) apresentaram efeitos significativos positivos, corroborando o proposto pela teoria, o que confirma as hipóteses 1a, 1b e 1c.

Para a análise da hipótese 1d, (o endividamento financeiro a valor contábil influencia negativamente e até certo ponto o risco de insolvência), foram analisados os resultados da variável endividamento financeiro a valor contábil (END_C_DF) e endividamento financeiro a valor contábil ao quadrado (END_C_DF_2). Tal análise levou em consideração a existência de uma relação quadrática entre o endividamento financeiro a valor contábil e o risco de insolvência na forma de U, de acordo com os pressupostos da teoria do *trade off*, que menciona o benefício fiscal advindo da redução do imposto devido ao pagamento de despesas financeiras, que incidem sobre o chamado passivo oneroso.

Por meio da análise, percebeu-se que a variável e seu termo quadrático apresentaram resultados significativos (p-valor=0,000). Conforme o esperado, a variável endividamento financeiro a valor contábil apresentou um coeficiente negativo, enquanto o termo quadrático da variável apresentou um coeficiente positivo, o que indica a relação de U esperada. Logo, os resultados confirmaram a hipótese 1d, o que mostra que o benefício fiscal proveniente do passivo oneroso influencia negativamente o risco de insolvência até certo ponto, tendo como ponto-ótimo o valor do endividamento financeiro (END_C_DF) igual a 0,4679, encontrado por meio da divisão do coeficiente da variável endividamento financeiro (END_C_DF), por duas vezes o coeficiente da variável quadrática de endividamento financeiro (END_C_DF_2).

Esses resultados indicam que o endividamento financeiro influencia negativamente e até certo ponto (END_C_DF = 0,4679) no risco de falência. Faz-se necessário ressaltar que o endividamento financeiro (END_C_DF) considera os valores sobre os quais há a incidência de

juros e, portanto, quanto maior seu valor, maior tenderá a ser o benefício fiscal proporcionado pelos juros pagos. De acordo com teoria do *trade-off*, o endividamento gera um benefício fiscal, devido à redução de impostos pagos decorrente da redução dada pelos juros pagos, mas os custos de falência também devem ser considerados.

Esse resultado pode ser considerado parcialmente consistente com o proposto por Miller (1974). De acordo com o autor, os custos de falência existem, mas podem ser proporcionalmente irrelevantes se comparados ao benefício fiscal do endividamento. Sendo assim, o benefício fiscal resultante do uso de capital de terceiros pode ser o responsável por compensar os custos de falência associados ao maior endividamento, diminuindo o risco de falência.

A significância do termo quadrático, no entanto, reforça a existência de um ponto ótimo de endividamento (BRITO; CORRAR; BATISTELLA, 2007; DURAND, 1959), bem como a existência do benefício fiscal, propostos pela teoria do *trade off* (MODIGLIANI; MILLER, 1963). Isso se deve ao fato de que, apesar de o benefício fiscal proporcionar uma diminuição no risco de insolvência, existe um ponto ótimo que maximiza os efeitos do benefício fiscal e minimiza os efeitos dos custos de falência (DEANGELO; MASULIS, 1980; MILLER, 1977; SCOTT, 1976).

Diferentemente do proposto por Miller (1977), o benefício fiscal compensa os custos de falência decorrentes do endividamento apenas até certo ponto. Isso indica que, apesar de o uso de capital de terceiros, especificamente de passivo oneroso, ser benéfico para as empresas, o uso não deve ser excessivo, pois de fato poderá levar a empresa a apresentar custos de falência também excessivos. Assim, as empresas devem usar o endividamento financeiro de forma moderada (DURAND, 1959), para minimizar seu risco de insolvência e maximizar seu valor.

Para analisar o efeito da pandemia de COVID-19 na relação entre estrutura de capital e risco de insolvência, adicionou-se ao modelo anterior as variáveis moderadoras obtidas por meio do produto entre as variáveis de estrutura de capital e a variável *dummy* que representou o período de pandemia (valores 1 para os quatro trimestres dos anos e 2020 e 2021; valor zero para os demais trimestres). A Tabela 16 mostra os resultados considerando as variáveis moderadoras.

Tabela 16 — modelo com variáveis moderadoras para o período de pandemia (continua).

Efeitos fixos				
Medida Repetida no nível 1				
Descrição	Variável	Coefficiente Beta	Erro Padrão	P-valor
	Trimestre	0.0010404	0.0094997	0.913

Tabela 16 — modelo com variáveis moderadoras para o período de pandemia (continua).

Hipóteses					
Hipótese	Variável	Coeficiente Beta	Erro Padrão	P-valor	Conclusão da Hipótese
1a	DIV_LIQ_C P	-2.22e-07	3.98e-08	0.000	Não confirma
1b	END_C_CP	5.223789	0.2656135	0.000	Confirma
1c	END_C_LP	5.242877	0.232654	0.000	Confirma
1d	END_C_DF	-3.738632	0.9464838	0.000	Confirma
	END_C_DF_2	4.14713	0.9199613	0.000	
2a	DIV_LIQ_C P_COVID	-2.33e-08	2.39e-08	0.331	Não confirma
2b	END_C_CP_ COVID	-0.9475935	0.1632067	0.000	Não confirma
2c	END_C_LP_ COVID	0.5679485	0.1476556	0.000	Confirma
2d	END_C_DF_ COVID	-0.4046438	0.4832239	0.402	Não confirma
2e	END_C_DF_2_ COVID	0.5216014	.4917393	0.289	Não confirma
-	DIV_LIQ_C P_TRIM	3.19e-09	1.33e-09	0.016	-
-	END_C_CP_ TRIM	0.0706315	0.0093459	0.000	-
-	END_C_LP_ TRIM	0.0391164	0.007883	0.000	-
-	END_C_DF_ TRIM	-0.0482987	0.0297745	0.105	-
-	END_C_DF_2_ TRIM	-0.0299663	0.0293486	0.307	-
-	COVID	0.0309511	0.1209358	0.798	-
Constante					
Item		Coeficiente Beta	Erro Padrão	P-valor	
_cons		-1.199877	0.4140726	0.004	
Efeitos aleatórios					
Item		Coeficiente	Erro Padrão	τ_{xxxx} / Erro-padrão	
τ_{u000} var(_cons) para Setor		0.1479642	0.1789982	0.826624	
τ_{u100} (var(data) para setor		0.0000192	0.0000268	0.7164179	
τ_{r000} var(_cons) para empresa		45.79878	3.763419	12.169461	
τ_{r100} (var(data) para empresa		0.0242931	0.001961	12.388118	
σ^2 var(Residual))		2.767231	0.0324842	85.186983	

Tabela 17 — modelo com variáveis moderadoras para o período de pandemia (conclusão).

Correlação Intraclasse		
Item	ICC	Erro Padrão
rhoempresa setor	0.9431943	0.0044634
rhoetor	0.0030374	0.0036772
Teste de Razão de Verossimilhança		
	Chi2 (2)	P-valor
LR-teste	27719.67	0.000

Fonte: Adaptado de Pereira (2021).

Para a análise da hipótese 2a (a pandemia de COVID-19 fortalece a influência positiva da dívida líquida de curto prazo no risco de insolvência), observou-se os resultados da variável moderadora dívida líquida de curto prazo X COVID-19 (DIV_LIQ_CP_COVID). O resultado da variável não foi significativo (p-valor=0,773) e negativo. Assim, a hipótese 2a do estudo não foi confirmada.

A análise da hipótese 2b (a pandemia de COVID-19 fortalece a influência positiva do endividamento de curto prazo no risco de insolvência) se deu por meio da observação da variável moderadora endividamento de curto prazo X COVID-19 (END_C_CP_COVID). De acordo com os resultados da Tabela 16, a variável apresentou resultado significativo (p-valor=0,000), porém negativo, indicando que a pandemia de COVID-19 enfraqueceu a influência positiva do endividamento de curto prazo no risco de insolvência. Logo, esse resultado não confirmou a hipótese 2b proposta.

Tal resultado pode ter sido decorrente do fato de que durante a pandemia de COVID-19, algumas empresas tiveram a paralização de suas atividades operacionais devido ao distanciamento. Ademais, face ao estado de calamidade a renegociação das dívidas (especialmente das de curto prazo), pode ter sido utilizada por diversas empresas, uma vez que a geração de caixa no período foi comprometida e também foram propostas regulações específicas para o período de pandemia.

A hipótese 2c (a pandemia de COVID-19 fortalece a influência positiva do endividamento de longo prazo no risco de insolvência) foi verificada por meio da análise da variável moderadora endividamento de longo prazo X COVID-19. Os resultados da variável foram significativos (p-valor=0,000) e com coeficiente positivo. Isso indica que a hipótese 2c foi confirmada. Sendo assim, no que se refere ao efeito do endividamento de longo prazo no risco de insolvência, a pandemia de COVID-19 de fato fortaleceu essa relação, fato que pode ser devido à necessidade do uso de capital de terceiros, que provavelmente foi negociado para

o longo prazo devido às possíveis dificuldades de geração de fluxo de caixa associadas à pandemia de COVID-19.

Para analisar as hipóteses 2d e 2e Hipótese 2d (A pandemia de COVID-19 enfraquece a influência negativa (até certo ponto) do endividamento financeiro a valor contábil no risco de insolvência; A pandemia de COVID-19 fortalece a influência positiva (a partir de certo ponto) do endividamento financeiro a valor contábil (termo quadrático) no risco de insolvência), observaram-se as variáveis moderadoras endividamento financeiro a valor contábil X COVID-19 e endividamento financeiro a valor contábil ao quadrado X COVID-19. Os resultados dessas variáveis não foram significativos (p-valor=0,402; p-valor=0,289), apesar dos coeficientes se apresentarem com sinais negativo e positivo, respectivamente. Esses resultados não confirmam as hipóteses 2d e 2e.

Assim, não foi possível confirmar que a pandemia enfraqueceu a relação inicialmente negativa entre o endividamento financeiro a valor contábil e o risco de insolvência, devido ao possível aumento do risco de insolvência decorrente, por exemplo, do risco de rolagem da dívida. Ademais, não foi possível confirmar que a pandemia fortaleceu a relação positiva, a partir de certo ponto, entre o endividamento financeiro a valor contábil e o risco de insolvência, proposta pela hipótese 2e.

5 CONSIDERAÇÕES FINAIS

O risco de insolvência pode ser considerado como um tema relevante para diversos agentes sejam eles tomadores de decisão, fornecedores, funcionários, entre outros. Uma empresa geralmente se torna insolvente quando não consegue gerar o fluxo de caixa suficiente para a realização de seus pagamentos conforme o vencimento (MATEUS, 2010). Existe também a chamada insolvência de saldo, que ocorre quando a empresa apresenta patrimônio líquido negativo, estando associado à estrutura de capital (MERTON, 1974).

Existem diversos fatores que podem influenciar o risco de insolvência, entre eles a estrutura de capital (TAO et al., 2020) e fatores do ambiente externo à empresa (HACKBARTH; MIAO; MORELLEC, 2006; TINOCO; WILSON, 2013). A estrutura de capital se refere ao modo como a empresa financia seus ativos, sendo por meio de capital próprio ou de terceiros (BRITO; CORRAR; BATISTELLA, 2007). Essa estrutura apresenta, portanto, o nível de endividamento das organizações, podendo ser crucial para o risco de insolvência.

Existem evidências de que fatores do ambiente macroeconômico também podem influenciar o risco de insolvência (HACKBARTH; MIAO; MORELLEC, 2006; TINOCO; WILSON, 2013). Um exemplo de fator externo foi a ocorrência da pandemia de COVID-19, que impactou a economia mundial e pode continuar gerando incertezas que podem influenciar o desempenho e o risco das organizações durante o período de crise (THE WORLD BANK, 2022).

Apesar da importância do risco de falência, a literatura se concentrou principalmente no desenvolvimento de modelos preditivos (BRYAN; DINESH FERNANDO; TRIPATHY, 2013) e a maioria dos estudos se concentrou em países desenvolvidos (OZ; SIMGA-MUGAN, 2018). Além do mencionado, nota-se também que apesar da existência de estudos sobre a estrutura de capital e o risco de insolvência, esses dois tópicos geralmente são pouco analisados conjuntamente, existindo também resultados mistos sobre a relação entre ambos.

Diante disso, para ampliar o conhecimento sobre os determinantes do risco de insolvência, o presente estudo teve como objetivo analisar o impacto da estrutura de capital no risco de insolvência de empresas brasileiras não financeiras de capital aberto negociadas na B3. Para as análises foram considerados dados trimestrais das demonstrações financeiras, obtidos por meio do Economática, considerando o período de janeiro de 2010 a dezembro de 2021. Para a análise do estudo, optou-se pelo uso da regressão multinível de dados em painel.

Por meio dos resultados, foi possível identificar que algumas das variáveis de estrutura de capital apresentaram resultados significativos, que confirmaram algumas das primeiras hipóteses propostas. Especificamente, as variáveis, endividamento de curto prazo (END_C_LP) e endividamento de longo prazo (END_C_LP) apresentaram efeitos significativos positivos, corroborando o proposto pela teoria do *trade-off*, que propõem que o uso de capital de terceiros aumenta os custos de falência e, portanto o risco de insolvência (MYERS, 2001). Esses resultados confirmaram as hipóteses 1b e 1c.

O resultado da variável dívida líquida de curto prazo (DIV_LIQ_CP), por sua vez foi significativo porém negativo, diferente do esperado. Logo a hipótese 1a do estudo não foi confirmada, indicando que a dívida líquida de curto prazo influenciou negativamente no risco de insolvência das empresas analisadas.

Ao analisar a variável endividamento financeiro a valor contábil (END_C_DF) e o termo quadrático dessa variável (END_C_DF_2), foi possível perceber que elas apresentaram resultados significativos, sendo a variável END_C_DF com coeficiente negativo e a variável END_C_DF_2 com coeficiente positivo. Esses resultados confirmaram a hipótese 1d do estudo, que, em que se esperava uma relação quadrática (em forma de U) entre o endividamento financeiro a valor contábil e o risco de insolvência. Tal resultado era esperado devido ao benefício fiscal advindo do passivo oneroso, que compensa, até certo ponto, os custos de falência associados ao uso da dívida. O Endividamento financeiro a valor contábil (END_C_DF) é composto pelo chamado passivo oneroso, ou seja, considera o capital de terceiros sobre o qual há juros.

Uma vez que a teoria do *trade-off* postula sobre o benefício fiscal proveniente do uso de capital de terceiros (MODIGLIANI; MILLER, 1963), esse resultado indica que o benefício fiscal proporcionado pelos juros pagos supera, até certo ponto os custos de falência, e portanto inicialmente influencia negativamente no risco de insolvência. Esse resultado pode ser considerado consistente pelo proposto pela teoria do *trade off*, que reconhece a existência dos custos de falência, mas defende a existência de um ponto ótimo para a estrutura de capital (BRITO; CORRAR; BATISTELLA, 2007; DURAND, 1959; MODIGLIANI; MILLER, 1963). Sendo assim, considera-se que o benefício fiscal diminui o risco de insolvência até certo ponto, mas caso o uso de capital de terceiros seja excessivo, os custos de falência superam esse benefício fiscal.

Ao considerar os efeitos da pandemia de COVID-19 na influência das variáveis de estrutura de capital no risco de falência, parte dos resultados não foram significativos ou não confirmaram as hipóteses do estudo. Especificamente, apenas a hipótese 2c (a pandemia de

COVID-19 fortalece a influência positiva do endividamento de longo prazo no risco de insolvência) foi confirmada.

Tal resultado sugere que a pandemia de COVID-19 fortaleceu a influência do endividamento de longo prazo no risco de insolvência. Esse resultado pode ser um indício de que diante da necessidade do uso de capital de terceiros, as empresas optaram principalmente por dívidas de longo prazo, devido às possíveis dificuldades decorrentes da pandemia de COVID-19. Pode ser também um efeito das possíveis renegociações das dívidas de curto prazo realizadas durante o período de pandemia.

Ao considerar o endividamento de curto prazo, apesar do resultado ter sido significativo, o coeficiente apresentou um resultado negativo, indicando que a pandemia de COVID-19 enfraqueceu a influência positiva do endividamento de curto prazo no risco de insolvência. Tal resultado pode ser devido à necessidade de renegociação das dívidas de curto prazo, face principalmente aos problemas relacionados à geração de caixa das empresas, uma vez que muitas necessitaram pausar suas atividades devido ao isolamento social realizado para conter o contágio do vírus.

Por meio dos resultados do estudo, percebe-se que este apresentou implicações acadêmicas, gerenciais, sociais e políticas. No que se refere às implicações acadêmicas, este estudo estendeu a literatura sobre o risco de insolvência, identificando fatores que podem influenciar nesse risco. Diferentemente da maioria dos estudos, que geralmente se preocupam com a elaboração de modelos preditivos, este estudo se concentrou nos efeitos da estrutura de capital e, por meio dos resultados, foi possível concluir que a teoria do *trade off* pode explicar a influência do endividamento no risco de insolvência. Assim, por meio deste estudo observou-se que existe um ponto ótimo que maximiza, os efeitos do benefício fiscal proveniente do passivo oneroso, e minimiza o risco de insolvência de empresas brasileiras não financeiras com ações negociadas na B3.

No que se refere às implicações gerenciais, por meio do estudo foi possível verificar que a estrutura de capital se faz relevante para o risco de insolvência das empresas analisadas. Tal conhecimento fornece aos tomadores de decisão indícios da existência de uma estrutura ótima de capital, que minimize o risco de falência. Sendo assim, os tomadores de decisão poderão adotar estratégias que favorecerão a saúde financeira das empresas, optando por uma estrutura de capital que não coloque em risco a empresa, o que favorecerá sua saúde financeira.

As implicações sociais do estudo referem-se aos benefícios gerados para a sociedade ao passo que uma empresa se mantém no mercado. Por meio do maior conhecimento sobre os fatores que influenciam no risco de insolvência, os tomadores de decisão terão conhecimento

que irá auxiliar na adoção de estratégias mais assertivas. Assim, todos os *stakeholders* ligados à organização irão se beneficiar, como investidores, funcionários e fornecedores, uma vez que a descontinuidade de uma empresa que chega à falência prejudica a todos os agentes envolvidos com a organização, fato que pode desencadear um efeito sistêmico.

Por fim, mencionam-se as implicações políticas. Conforme já mencionado, os estudos sobre risco de insolvência se concentraram principalmente nos modelos de previsão, mas entender os fatores que influenciam nesse risco também se faz relevante. Sendo assim, por meio desse conhecimento, políticas públicas mais adequadas podem ser formuladas, principalmente relacionadas à recuperação judicial, extrajudicial e falência, com o intuito de tornar principalmente a recuperação judicial e extrajudicial, mais eficientes e eficazes, para que as empresas consigam de fato se recuperar e tornarem-se competitivas no mercado.

Apesar da relevância do estudo, algumas limitações podem ser mencionadas. Primeiramente, foram considerados dados de empresas de capital aberto com ações negociadas na B3. Diante disso, a amostra não incluiu empresas que não estejam listadas na bolsa de valores, fato que pode ser considerado uma limitação. Para estudos futuros, sugere-se que amostras considerando diferentes empresas sejam realizados, como pequenas e médias empresas. Essa análise ajudará na ampliação do conhecimento sobre os determinantes do risco de insolvência, possibilitando uma visão mais abrangente dos efeitos da estrutura de capital no risco de insolvência de empresas de porte menor, sendo relevante analisar também a influência da pandemia nessas empresas.

Outra limitação do estudo foi o fato de serem consideradas apenas empresas brasileiras. Face às diferenças institucionais que podem influenciar na estrutura de capital e no risco de insolvência, faz-se relevante que estudos considerando empresas de outros países em desenvolvimento sejam consideradas, como empresas de países da América Latina. Tal análise poderá ampliar o conhecimento evidenciando se em diferentes contextos institucionais a influência da estrutura de capital apresenta resultados que corroboram ou contrapõem os encontrados no presente estudo, analisando também os efeitos da pandemia na relação e no risco de insolvência das empresas de diferentes países.

Por fim, menciona-se que no estudo não foram consideradas questões relacionadas à legislação referente à recuperação judicial, recuperação extrajudicial e falência. Para estudos futuros sugere-se que esse aspecto seja considerado, pois por meio de tal análise será possível identificar como a legislação vigente e as mudanças ocorridas ao longo dos anos influenciam na relação entre a estrutura de capital e risco de falência. Assim, uma nova perspectiva

relacionada aos fatores determinantes do risco de falência poderá ser analisada, também ampliando o conhecimento sobre a temática.

REFERÊNCIAS

- ALBANEZ, T.; VALLE, M. R. DO. Impactos da assimetria de informação na estrutura de capital de empresas brasileiras abertas Impacts of information asymmetry on the capital structure of Brazilian publicly-traded firms. **Revista Contabilidade & Finanças**, v. 20, n. 51, p. 6–27, 2009.
- ALBANEZ, T.; VALLE, M. R. DO; CORRAR, L. J. Fatores institucionais e assimetria informacional: influência na estrutura de capital de empresas brasileiras. **RAM. Revista de Administração Mackenzie**, v. 13, n. 2, p. 76–105, 2012.
- ALTMAN, E. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. **The Journal of Finance**, v. 23, n. 4, p. 589–609, 1968.
- ALTMAN, E.; BAIDYA, T.; DIAS, L. Previsão de problemas financeiros em empresas. **Revista de Administração de Empresas**, v. 19, n. 1, p. 17–28, 1979.
- ALVES, P.; FRANCISCO, P. The impact of institutional environment on the capital structure of firms during recent financial crises. **The Quarterly Review of Economics and Finance**, v. 57, p. 129–146, 2015.
- BAKER, M.; WURGLER, J. Market Timing and Capital Structure. **The Journal of Finance**, v. 57, n. 1, p. 1–32, 1 fev. 2002.
- BASTOS, D. D.; NAKAMURA, W. T. Determinantes da estrutura de capital das companhias abertas no Brasil, México e Chile no período 2001-2006. **Revista Contabilidade & Finanças**, v. 20, n. 50, p. 75–94, 2009.
- BASTOS, D. D.; NAKAMURA, W. T.; BASSO, L. F. C. Determinantes da estrutura de capital das companhias abertas na América Latina: um estudo empírico considerando fatores macroeconômicos e institucionais. **RAM. Revista de Administração Mackenzie**, v. 10, n. 6, p. 47–77, 2009.
- BITTENCOURT, W.; ALBUQUERQUE, P. Estrutura de capital: uma revisão bibliográfica das publicações dos periódicos nacionais. **Revista Contemporânea de Contabilidade**, v. 15, n. 34, p. 94–114, 2018.
- BRITO, G. A. S.; CORRAR, L. J.; BATISTELLA, F. D. Fatores determinantes da estrutura de capital das maiores empresas que atuam no Brasil. **Revista Contabilidade & Finanças**, v. 18, n. 43, p. 9–19, 2007.
- BRYAN, D.; FERNANDO, G.; TRIPATHY, A. Bankruptcy risk, productivity and firm strategy. **Review of Accounting and Finance**, v. 12, n. 4, p. 309–326, 2013.
- CARVALHO, P. V.; CURTO, J. D.; PRIMOR, R. Macroeconomic determinants of credit risk: Evidence from the Eurozone. **International Journal of Finance & Economics**, v. 27, n. 2, p. 2054–2072, 2022.
- CHAIA, A. J. **Modelos de gestão do risco de crédito e sua aplicabilidade ao mercado brasileiro**. 2003. Dissertação (Mestrado em Administração) – Faculdade de Economia,

Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2003.

CEPAL. **Anuario Estadístico de América Latina y el Caribe**. Disponível em: <<https://statistics.cepal.org/yearbook/2021/statistics.html?lang=es&theme=economicas>>. Acesso em: 13 set. 2022.

CNN. **Lucro das empresas de capital aberto cai 82% no segundo trimestre**. Disponível em: <<https://www.cnnbrasil.com.br/business/lucro-das-empresas-de-capital-aberto-cai-81-no-segundo-trimestre/>>. Acesso em: 26 nov. 2022.

DEANGELO, H.; MASULIS, R. W. Optimal capital structure under corporate and personal taxation. **Journal of Financial Economics**, v. 8, n. 1, 1980.

DEANGELO, H.; DEANGELO, L.; WHITED, T. M. Capital structure dynamics and transitory debt. **Journal of Financial Economics**, v. 99, n. 2, p. 235–261, 2011.

DURAND, D. **Costs of Debt and Equity Funds for Business: Trends and Problems of Measurement**. Conference on Research in Business Finance. **Anais...NBER**, 1952

DURAND, D. The Cost of Capital, Corporation Finance, and the Theory of Investment: Comment. **American Economic Review.**, v. 49, n. 4, p. 639–655, 1959.

DURAN, M. M.; STEPHEN, S. A. Internationalization and the capital structure of firms in emerging markets: Evidence from Latin America before and after the financial crisis. **Research In International Business and Finance**, v. 54, 2020.

FAHLENBRACH, R.; RAGETH, K.; STULZ, R. M. How Valuable Is Financial Flexibility when Revenue Stops? Evidence from the COVID-19 Crisis. **The Review of Financial Studies**, v. 34, n. 11, p. 5474–5521, 1 nov. 2021.

FÁVERO, L. P.; BELFIORE, P. **Manual de análise de dados: estatística e modelagem multivariada com Excel, SPSS e Stata**. 1. ed. Rio de Janeiro: Elsevier, 2017.

FREIRE, M. *et al.* Aderência às normas internacionais de contabilidade pelas empresas brasileiras. **Revista de Contabilidade e Organizações**, v. 6, n. 15, p. 3–22, 2012.

GUJARATI, D. N.; PORTER, D. C. **Econometria Básica**. 5ª ed. Brasil: McGraw Hill Brasi, 2011.

HACKBARTH, D.; MIAO, J.; MORELLEC, E. Capital structure, credit risk, and macroeconomic conditions. **Journal of Financial Economics**, v. 82, n. 3, p. 519–550, 2006.

HAIR, J. *et al.* **Análise Multivariada de Dados**. 6ª edição ed. São Paulo: Bookman Companhia Editora, 2006.

HARON, R. *et al.* The influence of firm, industry and concentrated ownership on dynamic capital structure decision in emerging market. **Journal of Asia Business Studies**, v. 15, n. 5, p. 689–709, 2021.

HOVAKIMIAN, A.; KAYHAN, A.; TITMAN, S. Are corporate default probabilities

consistent with the static trade-off theory? **The Review of Financial Studies**, v. 25, n. 2, p. 315–340, 2012.

HSIAO, C. Panel data analysis-advantages and challenges. **Test**, v. 16, n. 1, p. 1–22, 2007.

IWAKI, H. The effect of debt market imperfection on capital structure and investment: Evidence from the 2008 global financial crisis in Japan. **The Quarterly Review of Economics and Finance**, v. 74, p. 251–266, 2019.

JANG, B.-G.; RHEE, Y.; YOON, J. H. Business cycle and credit risk modeling with jump risks. **Journal of Empirical Finance**, v. 39, n. A, p. 15–36, 2016.

JÕEVEER, K. Firm, country and macroeconomic determinants of capital structure: Evidence from transition economies. **Journal of Comparative Economics**, v. 41, n. 1, p. 294–308, 2013.

LOTT, V. F. **Fatores Determinantes do Endividamento e Predição de Insolvência: Um Estudo Comparativo entre Brasil e Estados Unidos**. 2019. Dissertação (Mestrado em Administração) - Universidade Federal de Minas Gerais, Belo Horizonte, 2019.

LOTT, V. F.; TENENWURCEL, D. R.; CAMARGOS, M. A.. Determinants of Indebtedness of Brazilian Companies Listed in B3 With and Without Insolvency Risk. **Revista de Administração da UFSM**, v. 14, n. 1, p. 79–99, 2021.

LYUBOV, T.; HESHMATI, A. Impact of financial crises on the dynamics of capital structure: evidence from korean listed companies. **Singapore Economic Review**, p. 1–32, 2019.

MACHADO, M. A. V.; MEDEIROS, O. R.; EID JÚNIOR, W. Problemas na mensuração da estrutura de capital: evidências empíricas no Brasil. **Brazilian business review**, Vitória, v.7, n.1, p.24-47, jan./abr. 2010.

MERTON, R. C. On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. **The Journal of Finance**, v. 29, n. 2, p. 449–470, 3 jun. 1974.

MILLER, M. H. Debt and taxes. **Journal of Finance**, v. 32, n. 2, 1977.

MODIGLIANI, F., MILLER, M. The cost of capital, corporation finance and the theory of investment. **The American Economic Review**, p. 261–297, 1958.

MODIGLIANI, F.; MILLER, M. H. Corporate income taxes and the cost of capital: a correction. **The American Economic Review**, v. 53, n. 3, p. 433–443, 1963.

MYERS, S. The capital structure puzzle. **Journal of Finance**, v. 39, n. 3, p. 575–592, 1984.

MYERS, S. C. Capital structure. **The Journal of Economic Perspectives**, v. 15, n. 2, p. 81–102, 2001.

NENU, E.; VINTILĂ, G.; GHERGHINA, Ș. The Impact of Capital Structure on Risk and Firm Performance: Empirical Evidence for the Bucharest Stock Exchange Listed Companies. **International Journal of Financial Studies**, v. 6, n. 2, p. 41, 2018.

OZ, I. O.; SIMGA-MUGAN, C. Bankruptcy prediction models' generalizability: Evidence from emerging market economies. **Advances in Accounting**, v. 41, n. September 2017, p. 114–125, 2018.

PEREIRA, A. L. C. **Investimento em bens de capital: Determinantes e impacto no desempenho econômico-financeiro sob a moderação da governança corporativa**. 2021. Dissertação (Mestrado em Administração) – Universidade Federal de Lavras, Lavras, 2021.

PRADO, J. W. **Determinantes e implicações da estrutura de capital, da estrutura de propriedade e da governança corporativa: um modelo multiteórico de análise**. 2019. Tese (Doutorado em Administração) – Universidade Federal de Lavras, Lavras, 2019.

PRADO, J. *et al.* Multivariate analysis of credit risk and bankruptcy research data: a bibliometric study involving different knowledge fields (1968–2014). **Scientometrics, Budapeste**, v. 106, p. 1007–1029, 2016.

RICCA, L. T.; JUCÁ, M. N.; HADAD JUNIOR, E. Tax benefit and bankruptcy cost of debt. **Quarterly Review of Economics and Finance**, v. 81, p. 82–92, 2021.

SHI, Y.; LI, X. A bibliometric study on intelligent techniques of bankruptcy prediction for corporate firms. **Heliyon**, v. 5, n. 12, 2019b.

SILVA, A. I. S. **Um Modelo de Previsão de Insolvência Financeira: Aplicação ao Sector da Indústria Transformadora**. Dissertação (Mestrado em Economia) – Universidade da Beira Interior, Covilhã, 2010.

SILVA, E. D. S. *et al.* Capital structure of Brazil, Russia, India and China by economic crisis. **Revista de Administracao Mackenzie**, v. 17, n. 3, p. 105–131, 2016.

TAO, Q. *et al.* Does operating leverage increase firm's profitability and bankruptcy risk? Evidence from China's entry into WTO. **International Journal of Finance and Economics**, n. November, p. 1–17, 2020.

THE WORLD BANK. **A Economia Mundial Deve Expandir 4% em 2021; Implementação da Vacina e Investimentos são Essenciais para Sustentar a Recuperação**. 05 Jan. 2021. Disponível em: <<https://www.worldbank.org/pt/news/press-release/2021/01/05/global-economy-to-expand-by-4-percent-in-2021-vaccine-deployment-and-investment-key-to-sustaining-the-recovery>>. Acesso em: 30 mar. 2022.

THE WORLD BANK. **O crescimento global deve desacelerar até 2023, aumentando o risco de um declínio rápido e indesejado nas economias em desenvolvimento**. 11 Jan. 2022. Disponível em: <<https://www.worldbank.org/pt/news/press-release/2022/01/11/global-recovery-economics-debt-commodity-inequality>>. Acesso em: 30 mar. 2022.

TINOCO, M. H.; WILSON, N. Financial distress and bankruptcy prediction among listed companies using accounting, market and macroeconomic variables. **International Review of Financial Analysis**, v. 30, p. 394–419, 2013.

ZEITUN, R.; TEMIMI, A.; MIMOUNI, K. Do financial crises alter the dynamics of corporate capital structure? Evidence from GCC countries. **The Quarterly Review of Economics and Finance**, v. 63, p. 21–33, 2017.