Insper Instituto de Ensino e Pesquisa

Programa de Mestrado Profissional em Economia

PREVISÃO DE FALÊNCIAS CORPORATIVAS NO BRASIL: UMA ABORDAGEM MULTIPERÍODO COMPARANDO MÉTODOS TRADICIONAIS E APRENDIZADO DE MÁQUINA COM ÊNFASE NOS IMPACTOS ECONÔMICOS

Autor:

DIEGO FILIPPE ASSIS DE MIRANDA

Orientador: Prof. Paulo Sergio Oliveira Ribeiro

Coorientador: Prof. Magno Tairone de Freitas Severino

São Paulo
2024

DIEGO FILIPPE ASSIS DE MIRANDA

PREVISÃO DE FALÊNCIAS CORPORATIVAS NO BRASIL: UMA ABORDAGEM MULTIPERÍODO COMPARANDO MÉTODOS TRADICIONAIS E APRENDIZADO DE MÁQUINA COM ÊNFASE NOS IMPACTOS ECONÔMICOS

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper Instituto de Ensino e Pesquisa, como parte dos requisitos para a obtenção do título de Mestre em Economia.

Orientador: Prof. Dr. Paulo Sergio Oliveira Ribeiro

Co-orientador: Prof. Dr. Magno Tairone de Freitas Severino

São Paulo
2024

Miranda, Diego Filippe Assis de.

Previsão de falências corporativas no brasil: uma abordagem multiperíodo comparando métodos tradicionais e aprendizado de máquina com ênfase nos impactos econômicos / Diego Filippe Assis de Miranda; orientador: Prof. Paulo Sergio Oliveira Ribeiro – São Paulo: Insper, 2024.

Dissertação (Mestrado –Programa de Mestrado Profissional em Economia. Área de concentração: Finanças) –Insper Instituto de Ensino e Pesquisa.

1. Previsão de falências corporativas, Random Forest, CatBoost, Risco de Crédito. JEL: G33, G32 C52, M41,

FOLHA DE APROVAÇÃO

DIEGO FILIPPE ASSIS DE MIRANDA

PREVISÃO DE FALÊNCIAS CORPORATIVAS NO BRASIL: UMA ABORDAGEM MULTIPERÍODO COMPARANDO MÉTODOS TRADICIONAIS E APRENDIZADO DE MÁQUINA COM ÊNFASE NOS IMPACTOS ECONÔMICOS

Dissertação apresentada ao Programa de Mestrado Profissional em Economia do Insper Instituto de Ensino e Pesquisa, como requisito parcial para a obtenção do título de Mestre em Economia.

Área de concentração: Finanças

Aprovado em: 24/03/2024

Banca Examinadora

Prof. Paulo Sergio Oliveira Ribeiro - Orientador
Instituição: Insper – Instituto de Ensino e Pesquisa

Assinatura:_____

Prof. Dr. Bernardo De Oliveira Guerra Ricca
Instituição: Insper – Instituto de Ensino e Pesquisa

Assinatura:_____

Prof. Dr. Lucas Serrão Macoris
Instituição: FGV EAESP - Escola de Administração de Empresas de São Paulo

Assinatura:_____

AGRADECIMENTOS

Agradeço à vida por me proporcionar esta conquista e aos meus pais, Dirson e Cristina, por sempre acreditarem no poder da educação e da ciência. À minha esposa Mariana, pelo companheirismo, apoio e compreensão, especialmente ao longo destes intensos dois anos. A Renato Tyszler e Eduardo Paez, pelo incentivo e por acreditarem no meu potencial. A Luis Eduardo (Lecca), Dayana Xavier e Marcelo Hidalgo, por continuarem confiando em meu trabalho. A todos os amigos e familiares que estiveram presentes em meus pensamentos durante todos os finais de semana em que estive ausente. Aos meus professores Paulo Ribeiro e Magno Severino pela orientação assertiva e fundamental para a conclusão deste mestrado.

“Mestre não é quem sempre ensina, mas quem de repente aprende.”

(Grande Sertão: Veredas, João Guimarães Rosa)

RESUMO

Esta dissertação investiga a previsão de falências corporativas no Brasil por meio de uma abordagem multiperíodo, comparando métodos tradicionais, como o modelo Z-Score de Altman, a técnicas de aprendizado de máquina, incluindo Random Forest e CatBoost. A análise concentra-se na avaliação da eficácia desses modelos em diferentes horizontes temporais, com ênfase nas métricas estatísticas e nos impactos econômicos associados. Dentre os aspectos econômicos investigados, destacam-se o custo de classificação incorreta, a precificação de crédito e os efeitos no valor alavancado das empresas. Além disso, o trabalho destaca a importância de adaptar técnicas preditivas ao contexto brasileiro, que é caracterizado pela escassez de dados financeiros robustos e por desafios específicos ao ambiente econômico local. Os resultados evidenciam que os modelos de aprendizado de máquina não apenas superam os métodos tradicionais em termos de precisão preditiva, mas também apresentam potencial para aprimorar a tomada de decisões estratégicas, mitigar riscos financeiros e aumentar o valor das empresas no longo prazo.

LISTA DE FIGURAS

Figura 1: Framework de elaboração da curva ROC. 24
Figura 2: Curvas ROC-AUC e discriminação de empresas solventes de falidas. 25
Figura 3: Exemplo do impacto dos custos de classificação incorreta. 27
Figura 4: Evolução do número de empresas com evento default entre 1T08 e 2T24. ... 32
Figura 5: Distribuição dos pedidos de default dos dados da amostra e dados Serasa. ... 33
Figura 6: Distribuição acumulada de default das empresas da amostra. 33
Figura 7: Distribuição das variáveis preditoras. 44
Figura 8: Médias das variáveis preditoras ao longo do tempo. 45
Figura 9: Matriz de correlação das variáveis preditoras. 46
Figura 10: AUC dos modelos de previsão de probabilidades de falência. 57
Figura 11: Erro de calibração esperada. 59
Figura 12: CECI para 100 cenários da relação do custo do erro tipo I e tipo II 64
Figura 13: Valor alavancado da empresa para dois modelos de previsão de falência.... 73
Figura 14: Custo ponderado de capital para dois modelos de previsão de falência 73

LISTA DE TABELAS

Tabela 1: Medidas de avaliação de modelos de previsão de falência. 21

Tabela 2: Matriz de confusão. 22

Tabela 3: Variáveis preditores 36

Tabela 4: Estrutura de dados para cálculos de probabilidades condicionais *foward*..... 39

Tabela 5 Estatísticas comparativas entre empresas solventes e em default com teste t de significância 41

Tabela 6: Estatísticas das variáveis preditoras das empresas solventes. 42

Tabela 7: Estatísticas das variáveis preditoras das empresas que entraram em default. 43

Tabela 8: Erro de calibração esperado antes de depois da calibração das probabilidades..... 61

Tabela 9: Custo esperado de classificação incorreta para EMS e Random Forest para 4 trimestres. 63

Tabela 10: Análise comparativa de modelos considerando a política de concessão de crédito e a rentabilidade (conjunto de treinamento). 66

Tabela 11: Valor da empresa alavancada dado pelos modelos Random Forest e Altman (EMS) 71

Tabela 12: Distribuição da amostra por Porte e Natureza Jurídica. 82

Tabela 13: Distribuição da amostra por Porte e Setor. 83

SUMÁRIO

1 INTRODUÇÃO 11
1.1 Formulação do problema 13
1.2 Justificativa 13
1.3 Objetivos 14
1.3.1 Objetivo Geral 14
1.3.2 Objetivos específicos 14
1.4 Estrutura do trabalho 15
2 REVISÃO BIBLIOGRÁFICA 16
2.1 Evolução das Técnicas de Previsão de Falências 16
2.2 Medidas de Avaliação de Modelos de Previsão de Falências 20
3 DADOS E METODOLOGIA 31
3.1 Dados e definição de default 31
3.2 Variáveis explicativas 34
3.3 Modelo teórico e econométrico 36
3.4 Estatísticas descritivas 40
3.5 Divisão dos dados em Treino, Validação e Teste 46
3.6 Métodos de classificação 47
3.6.1 Altman Z-Score para mercados emergentes (EMS) 47
3.6.2 Regressão Logística 48
3.6.3 Random Forest 49
3.6.4 Light Gradient Boosting Machine 49
3.6.5 CatBoost 50
3.7 Critério de seleção dos melhores hiperparâmetros 52
3.8 Considerações sobre viés de sobrevivência nas amostras 53
4 RESULTADOS E DISCUSSÃO 55
4.1 Previsão das Probabilidades de Falência de Empresas Individuais 56
4.2 Calibração das Probabilidades 58
4.3 Custo esperado de classificação incorreta (CECI) 61
4.4 Rentabilidade de modelos de previsão de falências 65
4.5 Impacto dos modelos de previsão no valor da empresa alavancada 67
4.5.1 Cálculo do valor da empresa alavancada 67
4.5.2 Cálculo do valor da empresa desalavancada 68
4.5.3 Valor Presente dos benefícios fiscais da dívida 68
4.5.4 Cálculo do Custo da Dívida 68
4.5.5 Valor presente dos custos de dificuldades financeiras 69
4.5.6 Comparação do valor da empresa alavancada para diferentes modelos 70
5 CONSIDERAÇÕES FINAIS 74
REFERÊNCIAS BIBLIOGRÁFICAS 76
APÊNDICES 80
Apêndice A - AUC para diferentes horizontes de previsão (em trimestres) e métodos 80
Apêndice B - Tabela dos hiperparâmetros 80
Apêndice C – Distribuição da Amostra por Porte e Natureza Jurídica 82
ANEXOS 84
Anexo A - Teste McNemar 84
Anexo B - Cálculo dos ativos ponderados pelo risco (Risk-Weighted Assets, RWA) 85

1 INTRODUÇÃO

A previsão de falências corporativas constitui um elemento importante na mitigação de riscos financeiros e na otimização da alocação de recursos econômicos. Modelos preditivos robustos têm impacto direto não apenas no contexto corporativo, mas também em aspectos regulatórios, financeiros e econômicos mais amplos, influenciando decisões de crédito, estruturação de capital, precificação de empréstimos e estabilidade do sistema financeiro. Em contextos de mercados emergentes, como o Brasil, a precisão desses modelos torna-se ainda mais relevante devido à escassez de dados financeiros consistentes e às especificidades econômicas locais como taxas de juros elevadas e instabilidades econômicas mais recorrentes.

Este estudo avalia o desempenho de modelos de aprendizado de máquina na previsão de falências corporativas no Brasil em diferentes horizontes temporais, comparando-os aos métodos tradicionais. Os resultados demonstram que técnicas modernas como Random Forest apresentam superioridade estatística significativa em relação ao tradicional modelo Z-Score de Altman (EMS), especialmente na capacidade discriminatória entre empresas solventes e insolventes, medida pela área sob a curva ROC (AUC). Além disso, evidenciam-se ganhos econômicos, incluindo redução dos custos esperados de classificação incorreta, melhor ajuste na precificação do crédito (spreads mais aderentes ao risco real das empresas) e um aumento no valor econômico das empresas alavancadas, decorrente de estimativas mais confiáveis das probabilidades de default. Esses achados destacam o potencial de modelos avançados para aprimorar decisões estratégicas e a gestão de risco corporativo, motivando uma análise mais aprofundada sobre as categorias de modelos preditivos existentes e as suas limitações no contexto brasileiro.

Os modelos de previsão de risco financeiros existentes podem ser classificados em duas categorias principais: modelos estruturais e modelos de forma reduzida Sigrist & Leuenberger (2022). Modelos estruturais, como os propostos por Merton (1974) e Black & Scholes (1973), buscam modelar os mecanismos econômicos que levam uma empresa ao default. Em contrapartida, os modelos de forma reduzida, tradicionalmente explorados por autores como Altman (1968), Beaver (1966) e Sigrist & Leuenberger (2022), consideram que o risco de default depende de variáveis específicas das empresas e macroeconômicas, sem assumir um mecanismo específico de falência. Recentemente, abordagens de aprendizado de máquina, como Redes Neurais, Random Forest e técnicas

de boosting, têm sido utilizadas como modelos de forma reduzida, particularmente para previsões em períodos únicos. No entanto, conforme Sigrist & Leuenberger (2022), a aplicação de técnicas de aprendizado de máquina para modelar probabilidades de default em contextos multiperíodo ainda é limitada, representando uma lacuna importante na literatura.

Adicionalmente, a literatura internacional tem concentrado esforços em dados de mercados desenvolvidos, particularmente nos Estados Unidos, onde históricos de dados financeiros remontam à década de 1960, (Sigrist e Leuenberger ,2022, Barbosa e Altman et. al 2017). Isso contrasta com o contexto brasileiro, que enfrenta desafios relacionados à escassez de dados financeiros consistentes e à necessidade de adaptações metodológicas para modelos de previsão. No Brasil, a falta de estudos que explorem previsões multiperíodo e a escassez de dados confiáveis destacam a relevância de iniciativas que busquem gerar resultados robustos e adaptados à realidade local.

A crescente incidência de eventos de falência e recuperação judicial no Brasil reforça a necessidade de modelos preditivos robustos que não apenas melhorem a acurácia das previsões, mas também integrem análises econômicas para avaliar impactos em diferentes horizontes temporais. De acordo com dados da Serasa Experian, referentes ao primeiro semestre de 2024, houve um crescimento de 71% nos pedidos de recuperação judicial em comparação ao mesmo período de 2023, evidenciando a ampliação do risco financeiro entre as empresas brasileiras. Esse contexto, associado à escassez de estudos aplicados à realidade nacional, torna a investigação conduzida neste trabalho ainda mais premente, visando desenvolver modelos que sejam tecnicamente robustos e economicamente relevantes para a gestão de riscos corporativos.

Este trabalho busca preencher essa lacuna ao explorar como modelos de aprendizado de máquina podem ser aplicados ao contexto brasileiro, enfrentando desafios como a escassez de dados financeiros consistentes, as especificidades do mercado local e a necessidade de adaptações metodológicas. O objetivo é utilizar dados que permitam gerar resultados robustos e que sejam relevantes para apoiar decisões econômicas e financeiras em diferentes horizontes temporais. Além disso, investiga-se a relação entre a precisão preditiva e os impactos econômicos, como custos de classificação incorreta, valor alavancado das empresas com benefícios fiscais do endividamento e a precificação de crédito, ampliando a contribuição tanto para a teoria quanto para a prática na área.

Os resultados obtidos demonstraram que os modelos de aprendizado de máquina, como Random Forest, superaram métodos tradicionais, como o Z-Score de Altman, em precisão preditiva. Além disso, esses modelos mostraram impactos em três dimensões econômicas principais: na definição da estrutura ótima de capital, ao fornecer estimativas mais confiáveis das probabilidades de falência, permitindo ajustes mais eficientes na alavancagem empresarial; na precificação de crédito, ao calcular spreads mais ajustados aos perfis de risco, e na redução dos custos de classificação incorreta. Assim, conclui-se que as técnicas modernas de aprendizado de máquina oferecem contribuições relevantes para decisões estratégicas e a gestão de riscos corporativos considerando o mercado brasileiro.

1.1 Formulação do problema

A previsão de falências no Brasil enfrenta alguns desafios principais: a ausência de estudos que integrem análises econômicas e estatísticas, a falta de modelos multiperíodos e a necessidade de explorar modelos avançados de aprendizado de máquina. Além disso, a limitação de dados financeiros no Brasil dificulta a implementação de modelos adaptados ao contexto local. Diante disso, surgem as seguintes questões:

- Modelos modernos de aprendizado de máquina podem superar abordagens tradicionais, como o Z-Score de Altman, na previsão de falências no Brasil, especialmente em horizontes temporais mais longos?
- Como essas previsões podem ser utilizadas para mitigar os impactos econômicos das falências em empresas e credores?

1.2 Justificativa

Este trabalho apresenta uma justificativa em três dimensões inter-relacionadas. Primeiramente, no âmbito prático, o aumento expressivo nos pedidos de recuperação judicial em 2024, especialmente entre os diferentes portes empresariais, reforça a necessidade de ferramentas preditivas que antecipem crises financeiras e possibilitem decisões mais informadas. Segundo dados da Serasa Experian (2024), o volume de pedidos entre micro e pequenas empresas praticamente dobrou, passando de 376 pedidos no primeiro semestre de 2023 para 713 em 2024 — um crescimento absoluto de 337 casos e um crescimento relativo de 89,6%. As médias empresas também apresentaram um

aumento expressivo, com os pedidos passando de 155 para 207, representando um acréscimo de 52 casos, crescimento de 33,5%. Por sua vez, as grandes empresas experimentaram um crescimento ainda mais acentuado: de 62 para 94 pedidos, ou seja, um crescimento 51,6%.

Esses dados evidenciam que, embora o fenômeno atinja com mais intensidade as micro e pequenas empresas — em termos proporcionais e absolutos —, há uma tendência generalizada de elevação dos casos de recuperação judicial no Brasil. Esse cenário reforça a relevância de modelos preditivos robustos e adaptáveis a diferentes perfis empresariais, capazes de oferecer suporte analítico tanto para grandes corporações quanto para empresas de menor porte, tradicionalmente mais vulneráveis e com menor capacidade de gestão de riscos financeiros. Em segundo lugar, a contribuição acadêmica deste estudo reside na aplicação de modelos de aprendizado de máquina a dados brasileiros, explorando diferentes horizontes temporais e integrando métricas econômicas e estatísticas, uma abordagem pouco explorada no Brasil. Por fim, a relevância econômica se manifesta na capacidade de mitigar custos associados a previsões imprecisas, como inadimplência e desemprego gerados por eventos de falência corporativa, promovendo maior estabilidade financeira e sustentabilidade empresarial.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral deste estudo é avaliar o desempenho de modelos de aprendizado de máquina em comparação com métodos tradicionais na previsão de falências corporativas no Brasil, considerando diferentes horizontes temporais avaliados em aspectos econômicos e estatísticos.

1.3.2 Objetivos específicos

Para atingir o objetivo geral, este estudo estabelece os seguintes objetivos específicos: utilizar variáveis financeiras tradicionalmente relacionadas ao risco de falência no contexto de empresas brasileiras; aplicar metodologias avançadas como Random Forest e Catboost para previsões multiperíodo; comparar a performance desses modelos com o modelo de score de crédito para mercados emergentes (EMS) de Altman (2005); analisar

os impactos econômicos das previsões no valor alavancado das empresas; explorar a relação entre métricas estatísticas, como AUC, e métricas econômicas, como custo de classificação incorreta e valor da empresa alavancada; e avaliar o impacto desses modelos na precificação de crédito.

1.4 Estrutura do trabalho

Este estudo está organizado em cinco capítulos. O Capítulo 1 apresenta a introdução, contextualizando o tema, formulando o problema, justificando a pesquisa e delineando os objetivos e a estrutura do trabalho. O Capítulo 2 realiza uma revisão bibliográfica que abrange os avanços nos métodos de previsão de falências, destacando abordagens tradicionais e contemporâneas, além de revisar os métodos de avaliação e impacto econômicos de diferentes modelos de previsão de risco financeiro. O Capítulo 3 detalha os dados utilizados, os critérios de definição de falência e os métodos estatísticos e de aprendizado de máquina aplicados. O Capítulo 4 apresenta os resultados, discutindo o desempenho dos modelos e os impactos econômicos associados às previsões. Finalmente, o Capítulo 5 sintetiza os principais achados, discute suas implicações teóricas e práticas e propõe direções para pesquisas futuras.

2 REVISÃO BIBLIOGRÁFICA

Na literatura acadêmica de relevância, ainda há uma lacuna importante quanto à aplicação de métodos avançados de aprendizagem de máquina a empresas brasileiras. Essa deficiência parece estar associada às dificuldades de acesso a dados financeiros confiáveis, consequência do reduzido número de empresas de capital aberto no país — o que limitaria o acesso às informações contábeis e à centralização dos registros de falência de empresas de capital fechado.

Além disso, as abordagens utilizadas para medir o desempenho dos modelos de previsão de falência apresentam limitações que dificultam sua interpretação e a avaliação dos impactos econômicos. A revisão dessas limitações é aprofundada na seção 2.2, com destaque para os desafios inerentes às métricas atualmente utilizadas na literatura.

2.1 Evolução das Técnicas de Previsão de Falências

Historicamente, os métodos tradicionais de previsão de falências foram abordados por Beaver (1966), que utilizou Análise Discriminante Univariada. O autor analisou 14 índices financeiros e descobriu que a relação entre fluxo de caixa e dívida total, podia prever a situação financeira das empresas com um ano de antecedência. Posteriormente, Altman (1968) publicou um trabalho com Análise Discriminante Múltipla (MDA), introduzindo o modelo Z-Score. Este modelo evoluiu em relação ao anterior pois, ao invés de usar um modelo univariado utilizava múltiplos índices financeiras para discriminar entre empresas saudáveis e falidas. Ohlson (1980) aprimorou esta abordagem com o modelo O-Score, utilizando regressão logística para calcular probabilidades de falência de forma contínua.

Altman (1984), estendeu seu trabalho anterior mantendo a técnica de MDA, mas aplicando os modelos de previsão de falência em um contexto internacional, adaptando-os para considerar práticas contábeis e econômicas específicas de diferentes países, como o Brasil. Esta contribuição ajudou a resolver problemas de aplicabilidade dos modelos anteriores em diferentes ambientes econômicos e contábeis. Mensah (1984) acrescentou à literatura com a análise da estacionariedade dos modelos de previsão de falências, combinando os métodos de Análise Discriminante, Regressão Logística e destacando a necessidade de recalibração periódica para manter a precisão preditiva.

Apesar dos avanços das técnicas e da expansão de aplicabilidade dos modelos desde Beaver (1966) até Mensah (1984), esses estudos iniciais apresentaram algumas lacunas importantes, como a dependência excessiva de dados contábeis históricos, a falta de consideração de variáveis de mercado e a suposição de que as relações entre variáveis financeiras e a probabilidade de falência permanecem constantes ao longo do tempo. Essas limitações abriram caminho para o desenvolvimento de técnicas mais avançadas e flexíveis.

Com o avanço da tecnologia e o aumento da capacidade computacional, técnicas baseadas em aprendizagem de máquina e em dados de mercado começaram a ser exploradas para previsão de falências. Zhang et al. (1997) demonstraram que Redes Neurais Artificiais (ANNs) podem superar modelos tradicionais em precisão preditiva e robustez. No entanto, o estudo pode ter sido limitado pelo tamanho da amostra, o que pode afetar a generalização dos resultados para um contexto mais amplo. A solução para a limitação do tamanho da amostra pode ser encontrada no trabalho de Jones (2017) que aplicou técnicas de machine learning, como Gradient Boosting, para analisar dados de alta dimensionalidade, usando 91 variáveis preditoras, incluindo variáveis não tradicionais, como a remuneração de CEOs e a concentração acionária. Apesar dos avanços ao permitir a inclusão de um grande número de preditores e suas interações, a característica de "caixa-preta" nesses tipos de modelos dificultam sua aplicabilidade.

A análise de componentes principais (PCA) utilizada por Canbas et al. (2005) pode ajudar a abordar a questão das "caixas pretas" ao simplificar a complexidade dos dados financeiros em fatores principais que são mais fáceis de interpretar. Este método reduz a dimensionalidade dos dados, destacando os componentes mais importantes que explicam a variabilidade nos dados financeiros das empresas. O trabalho de Jardin (2016), também colaborou para melhorar a interpretabilidade ao introduzir uma técnica de classificação em duas etapas, combinando Mapas Auto-organizáveis de Kohonen (SOM's) e técnicas de Ensemble. As técnicas SOM's são usadas para identificar e quantizar perfis financeiros distintos de empresas. Isso permite a criação de modelos específicos para diferentes tipos de perfis financeiros, reconhecendo que as empresas podem seguir diferentes caminhos para a falência. A abordagem de técnicas híbridas é reforçada no trabalho de Kumar et al. (2007) que revisaram técnicas estatísticas para previsão de falências em bancos e empresas não financeiras. Esse estudo destacou a eficácia de métodos como a análise discriminante linear (LDA), regressão logística, redes neurais artificiais (ANNs) e

máquinas de vetores de suporte (SVMs). A revisão concluiu que técnicas combinadas de métodos estatísticos, apresentam maior precisão preditiva.

Além de contribuir com a hipótese de que combinações de técnicas são mais eficientes que apenas um modelo específico, o trabalho de Kumar et al. (2007) aplicou as técnicas de previsão em setores diferentes considerando empresas financeiras e não financeiras. Outro trabalho que colabora com os aspectos de diversificação de contextos é o de Zuppini et al. (2021) que aplicaram técnicas de machine learning no contexto de precificação de debêntures ilíquidas, combinando técnicas de clustering e modelos supervisionados para prever a variação do spread dessas debêntures. Os experimentos demonstraram que a combinação de técnicas, mesmo em contextos distintos aos de previsão de falências, resultou em estimativas mais precisas e em um aumento significativo no coeficiente de determinação (R^2) em comparação com o uso de cada técnica separadamente.

Apesar dos trabalhos de Canbas et al. (2005) resolverem os problemas de interpretabilidade com PCA, Jardin (2016) com SOM's e Kumar et al. (2007) junto com o trabalho de Zuppini et al. (2021) reforçarem que a combinação de técnicas pode ser mais precisa, é importante reconhecer que a aplicação de técnicas em diferentes contextos pode exigir adaptações significativas, visto que as diferenças econômicas e contábeis entre contextos podem ser complexas de modelar. Além disso, alguns desses trabalhos não estenderam os testes ao longo de períodos mais longos para fornecer uma validação mais abrangente dos modelos.

Para resolver a lacuna de aplicação em diferentes contextos, os trabalhos de Almeida et al. (2008) e Hillegeist et al. (2004) apresentaram algumas soluções. Os primeiros propuseram um modelo de estrutura a termo para precificar títulos sujeitos a *default*, extraindo probabilidades de *default* a partir de dados de títulos soberanos brasileiros, demonstrando a aplicabilidade de modelos dinâmicos em contextos diferentes das previsões de falências. Hillegeist et al. (2004) também abordaram outro contexto ao introduzirem o modelo BSM-Prob, que integra variáveis de mercado utilizando o modelo de precificação de opções de Black-Scholes-Merton, oferecendo uma medida mais ampla ao incorporar informações de mercado além dos dados contábeis tradicionais. No entanto, no modelo BSM de Hillegeist et al. (2004), há a suposição de que as taxas de perda (a porcentagem do valor do título que não é recuperada em caso de *default*) permanecem constantes ao longo do tempo, o que pode limitar a precisão do modelo. Esta suposição

pode não refletir as variações reais das taxas de recuperação em diferentes períodos e condições de mercado.

Nesse sentido, Jardin (2015) avançou na utilização de processos de falência terminal, que melhoram previsões de longo prazo ao considerar a evolução financeira das empresas ao longo do tempo. Nesse trabalho os autores concluem que o uso do processo de falência terminal também reduziu a quantidade de dados necessários para três anos, facilitando a aplicação prática e tornando o modelo acessível para empresas mais jovens ou aquelas com dados históricos limitado. Trabalhos realizados por Sigrist e Leuenberger (2022) utilizando técnicas de machine learning para prever o risco de *default* corporativo também apresentaram soluções para abordagens ao longo de múltiplos períodos, aplicando métodos como Tree-boosting e modelos de fragilidade para capturar correlações residuais e melhorar previsões de cauda. Esses métodos demonstraram uma precisão superior nas previsões em relação aos modelos lineares tradicionais, especialmente em horizontes de previsão mais longos.

Outros trabalhos que ajudaram na resolução dos problemas de contexto são os artigos de Rajan et al. (2015) e Zeidan et al. (2023). Os primeiros utilizaram uma abordagem empírica, combinando análise de séries temporais e modelos econométricos, para investigar o impacto da securitização em inadimplências. Eles constataram que os modelos tradicionais falham em capturar as mudanças comportamentais dos agentes econômicos induzidas pela securitização. Zeidan et al. (2023) introduziram o sistema de classificação de crédito de sustentabilidade generalizado, incorporando variáveis ESG (ambientais, sociais e de governança) para melhorar a análise de risco corporativo. Porém, a falta de consenso ou padronização global sobre quais fatores ESG devem ser considerados e como devem ser ponderados na avaliação de risco pode levar a inconsistências na interpretação e na aplicação dos critérios ESG.

Comparando a evolução das técnicas de previsão, nota-se que métodos tradicionais, como o Z-Score e o O-Score, são fáceis de implementar e interpretar, mas limitados em capturar dinâmicas complexas e mudanças econômicas. Em contraste, técnicas modernas de machine learning, como ANNs e Gradient Boosting, oferecem maior precisão e flexibilidade, lidando com grandes volumes de dados e relações não lineares. No entanto, essas técnicas exigem grandes conjuntos de dados para treinamento e são difíceis de interpretar. Alguns estudos mostram que a combinação de técnicas melhora a precisão das previsões, apesar de sua implementação ser mais complexa.

Para pesquisas futuras, os autores sugerem integrar dados financeiros e não financeiros das empresas e há ainda espaço para utilização de informações que abordam padrões ESG. A combinação de técnicas de machine learning, como clustering, Gradient Boosting e redes neurais, é recomendada para lidar com a alta dimensionalidade e complexidade dos dados financeiros, aumentando a precisão preditiva (Jones, 2017; Sigrist e Leuenberger, 2022). Estudos adicionais de validação são essenciais para garantir a robustez e confiabilidade dos modelos. A adaptação contínua dos modelos a diferentes condições econômicas e comportamentais deve ser considerada para manter sua relevância e precisão ao longo do tempo (Mensah, 1984; Rajan, 2015).

Essa dissertação foca na solução de alguns desafios identificados na previsão de falências corporativas, integrando abordagens avançadas de aprendizagem de máquina em comparação com o modelo de Altman (1983) para mercados emergentes. O modelo proposto será aplicado em diferentes horizontes temporais contemplando assim contextos econômicos distintos, garantindo uma análise adaptável às mudanças no ambiente financeiro.

2.2 Medidas de Avaliação de Modelos de Previsão de Falências

A revisão apresentada no tópico anterior avaliou os modelos de previsão na perspectiva do método; esta seção aborda como essas metodologias são avaliadas do ponto de vista estatístico e algumas abordagens para medir impactos econômico-financeiros entre os agentes envolvidos nas decisões de crédito.

As avaliações de eficiência econômica dos modelos são importantes, pois auxiliam as empresas tomadoras e os credores a identificarem com mais confiabilidade os riscos de insolvência, mitigando problemas de seleção adversa, como descrito por Akerlof (1970) em seu estudo sobre mercados com informações assimétricas, onde os desequilíbrios informacionais podem levar à exclusão de bons tomadores de crédito e ao aumento do risco no sistema.

A tabela 1 apresenta uma visão geral das métricas abordadas neste trabalho para a avaliação de modelos de previsão de falências. A grande maioria dos estudos nessa área usam como principal métrica a AUC (*area under the curve ROC)* que é útil para descrever a capacidade discriminatória dos modelos, enquanto outras medidas menos usuais integram aspectos econômicos e financeiros, como o CECI (custo esperado de

classificação incorreta), o Spread de Crédito para precificação de dívidas e o impacto das previsões de probabilidades sobre o valor da empresa alavancada. O conjunto dessas quatro perspectivas permite mensurar não apenas a precisão discriminatória, mas também os impactos econômicos das decisões tomadas com base nos modelos. Nos parágrafos subsequentes, é explorado mais detalhadamente cada uma dessas métricas, suas aplicações e implicações no contexto da análise de crédito e previsão de insolvência.

Tabela 1: Medidas de avaliação de modelos de previsão de falência.

Métrica	Tipo	Descrição	Vantagens	Principais referências na Literatura
Área Sob a Curva (AUC)	Discriminatória	Mede a capacidade do modelo de distinguir entre empresas solventes e falidas em diferentes limiares de decisão.	Permite comparação relativas entre modelos.	James e Tibshirani et al. (2023), Sigrist e Leuenberger (2022)
Custo Esperado de Classificação Incorreta (CECI)	Econômico-Financeira	Calcula o custo esperado dos erros de classificação, ponderando probabilidades a priori e custos relativos de erros tipo I e tipo II.	Integra custos econômicos na avaliação do modelo.	Frydman, Altman e Kao (1985); Jardin (2015, 2016)
Precificação de Crédito	Econômico-Financeira	Define o spread mínimo necessário com base no risco de inadimplência, considerando perda em caso de default (LGD) e taxa de juros mínima exigida pelos credores	Traduz previsões para decisões práticas de crédito.	Blöchlinger e Leippold (2006); Bauer and Agarwal (2014); Lohmann e Ohliger (2022);
Valor da Empresa Alavancada	Econômico-Estratégica	Avalia o impacto das previsões no valor econômico da empresa, incorporando benefícios fiscais da dívida e custos esperados de falência.	Conecta previsões a decisões estratégicas de estrutura de capital e maximização de valor.	Modigliani e Miller (1958); Altman (2006); Rajan et al. (2015)

A acurácia é a medida mais intuitiva para avaliar esses modelos, ela é calculada com o total de acertos sobre o total de previsões do modelo. Embora seja uma métrica intuitiva

tem várias limitações para modelos de classificação de falência. Ela não diferencia entre os tipos de erros de classificação, como classificar uma empresa que irá falir como solvente ou classificar uma empresa que terá dificuldades financeiras como solvente.

$$Acurácia = \frac{Quantidade\ de\ classificações\ corretas}{Total\ de\ observações}$$

Para superar as limitações da acurácia os métodos de classificação são avaliados em termos de outras taxas, derivadas da matriz de confusão, James et al. (2023), que compara os resultados do modelo com os dados reais. A forma como estruturar essas medidas na matriz de confusão está representada na Tabela 2:

Tabela 2: Matriz de confusão.

Matriz de confusão	Predito: (1 | Falência)	Predito: (0 | Solvente)	
Real: (1 | Falência)	Verdadeiro Positivo (VP)	Falso Negativo (FN)	Total empresas realmente falidas (VP + FN)
Real: (0 | Solvente)	Falso Positivo (FP)	Verdadeiro Negativo (VN)	Total empresas realmente solvente (FP + VN)
	Total predito como Falidas (VP + FP)	Total predito como solvente (VN + FN)	Total empresas na amostra (VP+FN+FP+VN)

A matriz de confusão é a base cálculo de taxas de erro/acerto de modelos de classificação. Fonte: AUTOR (2024), adaptado de James e Tibshirani et al. (2023)

As medidas mais importantes no contexto de previsão de falências oriundas dessa matriz são: (i) o erro tipo II que ocorre quando uma empresa solvente é incorretamente classificada como falida. Pode ser chamada de taxa de falsos positivos (TFP), que mede a proporção de empresas solventes classificadas erroneamente como falidas em relação ao total real de solventes.

$$Erro\ tipo\ II = (TFP) = \frac{FP}{FP + VN} = \frac{Solventes\ prevista\ como\ falidas}{Total\ real\ de\ solventes}$$

(ii) o erro tipo I ocorre quando uma empresa falida é incorretamente classificada como solvente. Pode ser chamada de taxa de falsos negativos (TFN), que mede a proporção de empresas falidas classificadas erroneamente como solventes em relação ao total real de falidas.

$$Erro\ tipo\ I = (TFN) = \frac{FN}{FN + VP} = \frac{Falida\ prevista\ como\ solvente}{Total\ real\ de\ falidas}$$

(iii) a sensibilidade, ou taxa de verdadeiros positivos (TVP), mede a proporção de empresas corretamente classificadas como falidas em relação ao total real de falidas;

$$Sensibilidade = (TVP) = \frac{VP}{VP + FN} = \frac{Falidas\ \ prevista\ como\ falidas}{Total\ real\ de\ falidas}$$

(iv) a precisão, mede a proporção de empresas corretamente classificadas como falidas em relação ao total de empresas previstas como falidas

$$Precisão = \frac{VP}{VP + FP} = \frac{Falidas\ \ prevista\ como\ falidas}{Total\ predito\ como\ falidas}$$

(v) a especificidade mede a proporção de empresas solventes corretamente classificadas como solventes em relação ao total real de solventes.

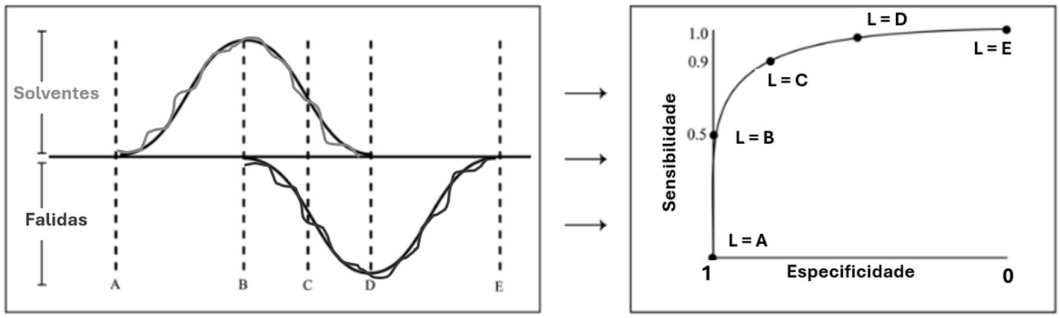
$$Especificidade = \frac{VN}{VN + FP} = \frac{Solventes\ prevista\ como\ solventes}{Total\ real\ de\ solventes}$$

Para verificar essas taxas de erros, é preciso determinar um limiar ou corte de classificação, ou seja, um valor que determina o ponto a partir do qual o modelo decide se uma empresa pertence à classe "falidas" ou "solventes".

A partir dessas três características (sensibilidade (TVP), especificidade e limiar de decisão), é traçada a curva ROC (Receiver Operating Characteristic**),** um padrão na avaliação de modelos preditivos.

Observa-se na Figura 1 à direita, uma curva ROC típica com a taxa de verdadeiro positivo (sensibilidade) no eixo vertical contra a Especificidade (com a escala invertida de 1 a 0) no eixo horizontal para uma faixa de valores de corte. Ao deslocar o limiar *L* de *A* para *E* na figura à esquerda, os pontos correspondentes ao longo da curva ROC são gerados. Assim, a curva ROC permite visualizar o desempenho do modelo ao longo de vários limiares, e sua inclinação revela o equilíbrio entre verdadeiros positivos e falsos positivos.

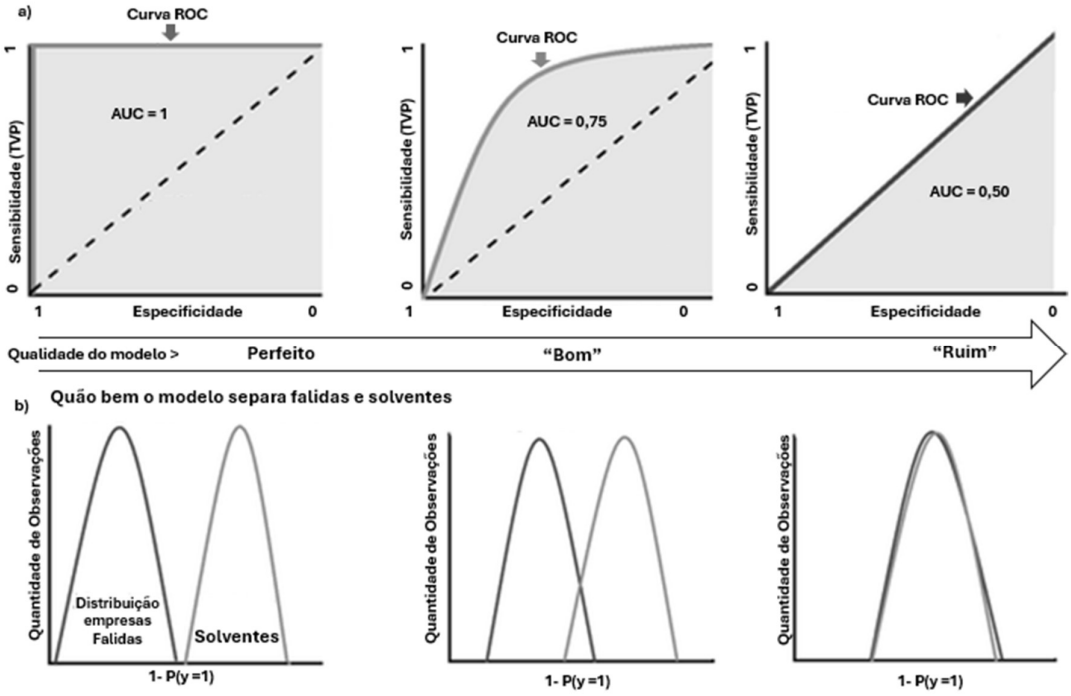
Figura 1: Framework de elaboração da curva ROC.



Ao lado esquerdo diferentes limiares de decisão (A até E) para probabilidades de falências atribuídas a empresas, variando de 0 a 1 e à direita a curva ROC com as combinações de sensibilidade e especificidade para cada limiar de decisão escolhido para classificar solventes de falidas. Fonte: AUTOR (2024), adaptado de Data Science Central (2019);

A métrica amplamente usada para avaliar cada curva é a Área sobre a curva ROC (AUC). Quanto mais próximo de um for essa área, mais precisamente o modelo discrimina empresas solventes de insolventes. Na Figura 2 observa-se um modelo perfeito em que a AUC é 1, passando por um modelo intermediário ("bom") com AUC igual a 0,75 até um modelo de classificação aleatório com AUC igual à 0,5. Nos gráficos de distribuição abaixo de cada curva ROC há uma ilustração de como cada um dos modelos consegue discriminar as empresas "boas" das "ruins", atribuindo mais ou menos riscos para cada uma delas.

Figura 2: Curvas ROC-AUC e discriminação de empresas solventes de falidas.



Fonte: AUTOR (2024), adaptado de Data Science Central (2019). Nos gráficos de probabilidades na parte inferior da figura o eixo X representa a probabilidade de uma empresa ser solvente.

No entanto, a AUC não incorpora a assimetria de custos relacionados aos erros tipo I e tipo II de modelos de classificação para falências empresariais. Do ponto de vista econômico-financeiro o erro tipo I geralmente apresenta um custo significativamente mais alto que o erro tipo II, pois o primeiro é diretamente impactado pelo custo do empréstimo cedido à uma empresa incorretamente classificada como solvente e que na verdade que ficará insolvente, enquanto o segundo tipo de erro tem apenas o custo de oportunidade do credor que deixou de ceder um empréstimo à uma empresa solvente.

Reconhecendo que erros do Tipo II e do Tipo I têm consequências financeiras distintas, a avaliação dos modelos evoluiu para incorporar os impactos da assimetria dos custos de classificação incorreta.

Frydman, Altman e Kao (1985) abordaram a questão da sensibilidade aos custos em modelos de classificação, propondo uma métrica de avaliação que combina os custos relativos dos erros do Tipo I e Tipo II com as probabilidades a priori de falência e não falência. Eles introduziram o conceito de Custo Esperado de Classificação Incorreta (CECI), calculado como:

$$Custo\ esperado\ de\ classificação\ incorreta\ (CECI) = \frac{c_1 . p_1 . e_1}{n_1} + \frac{c_2 . p_2 . e_2}{n_2}$$

Onde:

- c_1 e c_2: custos dos erros Tipo I e Tipo II, respectivamente;
- e_1 e e_2: taxas de erros Tipo I e Tipo II, respectivamente;
- n_1 e n_2: número de empresas solventes e falidas na amostra, respectivamente;
- p_1 e p_2: probabilidades a priori de não falência e de falência, respectivamente

Jardin (2015, 2016) propôs um framework para ajustar o *cut-off* (limiar de decisão) em modelos de classificação, permitindo que as decisões sejam adaptadas aos custos relativos dos erros do Tipo I e Tipo II, bem como às probabilidades a priori de falência e não falência. O ponto de corte ideal é determinado pelo balanceamento dos custos marginais: o custo incremental de reduzir erros Tipo I deve ser igual ao custo incremental de aumentar erros Tipo II. Quando a amostra é desbalanceada, ou seja, o número de empresas solventes (n_1) é diferente do número de empresas falidas (n_1), o cálculo do *cut-off* ideal é ajustado para levar em conta esse desequilíbrio e é dado por:

$$Cut - off\ ideal = \frac{c_2}{c_1 + c_2}.$$

Nessa relação, o *cut-off* reflete a proporção do custo do erro Tipo II (c_2) em relação ao custo total dos erros $(c_1 + c_2)$, garantindo que a decisão seja tomada de forma a minimizar o Custo Esperado de Classificação Incorreta (CECI). Essa abordagem fundamenta-se na premissa de que decisões baseadas em modelos preditivos devem priorizar os objetivos econômicos do credor, considerando a assimetria nos custos de classificação incorreta.

Na Figura 3 observa-se dois cenários hipotéticos para exemplificar os impactos do CECI em um contexto de previsão de falência. Nesse exemplo uma empresa A deseja captar \$5.000,0 a uma taxa de juros de 3%. Se o credor rejeitar o crédito e a empresa não falir ele terá o custo de oportunidade de \$150,0, se o credor ceder o crédito e a empresa falir o custo será a perda de 50% do valor da dívida, \$2.500,0.

No cenário 1 essas diferenças de custos não são consideradas. Para os dois cenários considerados há três níveis de probabilidades à priori e taxas de erro tipo 1 (tx_1) e erro tipo 2 (tx_2). Na coluna A está demonstrando uma taxa de erro tipo I e tipo II de 1%, ou

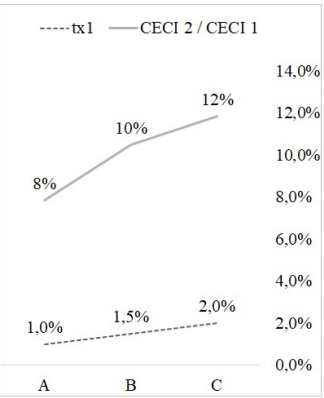
seja, o modelo classificou apenas 1% das empresas que não faliram como insolventes e 1% das empresas falidas como solventes. A probabilidade à priori de falência dessa economia, ou seja, a probabilidade calculada pelos dados históricos reais é de 0,5% na hipótese A, 1% na hipótese B e 1,5% na hipótese C. Assim, no primeiro cenário e na hipótese A, o custo esperado de classificação incorreta dado pela equação do *cut-off* é de $1,5. No cenário 2 hipótese A, onde é considerado a diferença de custos entre os tipos de erros o CECI passa a ser $1,6, ou 8% a mais que o custo do cenário 1, quando as probabilidades de falência à priori aumentam em 0,5% a diferença de custos aumentam para 10% (hipótese B) e em seguida para 12% (hipótese C).

Figura 3: Exemplo do impacto dos custos de classificação incorreta.

Empresa A	**Valor**
Dívida ($)	5.000
juros	3%
c1 (% dívida)	3%
c1 ($)	150,0
c2 (custo/oportunidade)	50%
c2 ($)	2.500,0
c2/c1	16,7

Legenda (Cenário 1 e 2)
c1 = custo do erro tipo 1
p1 = prob. a priori de falência
tx1 = taxa de erro tipo 1
c2 = custo do erro tipo 2
p2 = prob. a priori de não falência
tx2 = taxa de erro tipo 2

		A	**B**	**C**
CECI - Cenário 1		**1,5**	**2,2**	**3,0**
Cenário em que c1 = c2	*c1*	*150,0*	*150,0*	*150,0*
	p1	99,5%	99,0%	98,5%
	tx1	1,0%	1,5%	2,0%
	c2	*150,0*	*150,0*	*150,0*
	p2	0,5%	1,0%	1,5%
	tx2	1,0%	1,0%	1,0%
CECI - Cenário 2		**1,6**	**2,5**	**3,3**
Cenário em que c1 = (16,7 x c2)	*c1*	*150*	*150*	*150*
	p1	99,5%	99,0%	98,5%
	tx1	1,0%	1,5%	2,0%
	c2	*2.500,0*	*2.500,0*	*2.500,0*
	p2	0,5%	1,0%	1,5%
	tx2	1,0%	1,0%	1,0%



Fonte: AUTOR (2024)

Esse exemplo ilustra, portanto, a limitação de avaliar modelos de previsão de falência apenas pelas medidas mais tradicionais como a curva ROC e a AUC ao demonstrar o impacto relevantes dos custos dos tipos de erros.

Outra avaliação complementar proposta por Blöchlinger e Leippold (2006) expande a análise das curvas ROC para incluir o cálculo do valor presente líquido (NPV, na sigla em inglês) como critério para decisões de crédito ótimas. Essa metodologia conecta a característica discriminatória dos modelos de classificação de risco à precificação de spreads de crédito, permitindo que credores otimizem seus lucros ao ajustar as condições dos empréstimos com base nas probabilidades de inadimplência e no perfil de risco de cada cliente.

Essa metodologia usa o NPV para definir regras de aceitação e precificação de crédito. O NPV por unidade de crédito, ajustado para um horizonte de tempo de um período, foi definido pela fórmula:

$$NPV\ (t) = \frac{P(Y = 1|S = t).(1 - LGD) +\ P(Y = 0|S = t).(1 + R + C)}{1 + d} - 1$$

Onde;

- $P(Y = 1|S = t)$: probabilidade condicional de inadimplência dado o escore t;
- $P(Y = 0|S = t)$: probabilidade condicional de não inadimplência dado t;
- LGD = perda em caso de inadimplência (*Loss Given Default)*;
- $R =$ taxa de juros paga ao vencimento (spread de risco);
- C = valor estratégico ou benefício relacional esperado do cliente;
- $d =$ taxa de desconto ajustada ao risco.

A partir dessa formulação e desconsiderando o valor estratégico C, a condição $NPV\ (t) \geq 0$ pode ser reorganizada para derivar uma expressão para o spread mínimo requerido pelo credor dado por $R\ (t)$. Isso é obtido ao impor que o retorno líquido esperado do empréstimo, ajustado pela probabilidade de inadimplência e pela perda dada a inadimplência (LGD), seja não negativo, resultando na equação que estabelece um spreado mínimo dado por:

$$R\ (t) = \frac{P(Y = 1|S = t)}{P(Y = 0|S = t)}.LGD + k.$$

Onde k representa a taxa para cobrir o retorno mínimo exigido pelos acionistas ou para refletir condições de mercado específicas.

Essa relação mostra que o spread necessário é proporcional ao risco de inadimplência ajustado pelo LGD e pelas probabilidades condicionais de default $P(Y = 1|S = t)$ e não-default $P(Y = 0|S = t)$.

A avaliação dos spreads mínimos de crédito gerados por diferentes modelos de classificação é uma ferramenta para identificar qual modelo oferece melhores resultados econômicos e financeiros. Ao comparar dois modelos de previsão, cada observação de empresa gera um spread de crédito distinto. O modelo que atribuir o menor spread para uma empresa solvente será preferido, enquanto empresas com alto risco de falência podem ter seu crédito recusado se o spread calculado exceder um limite aceitável. Essa

análise permite avaliar o impacto econômico na atração de empresas solventes e na exclusão de empresas de alto risco do portfólio de crédito.

Além dessas métricas de avaliação, a análise do Valor da Empresa Alavancada (V^L) permite compreender como os modelos de previsão de falência influenciam decisões financeiras estratégicas. O V^L incorpora o equilíbrio entre os benefícios fiscais da dívida e os custos esperados de falência, conceitos derivados da Proposição II de Modigliani e Miller (1958), que demonstrou como a estrutura de capital pode impactar o custo de capital e o valor da empresa ao incluir os benefícios fiscais da dívida. Essa formulação convergiu para a Teoria do *Trade-off*, segundo a qual as empresas buscam uma estrutura de capital ótima que maximize seu valor, equilibrando os benefícios fiscais do endividamento com os custos crescentes de falência DeMarzo (2010).

A partir dessa abordagem, os impactos de diferentes modelos de previsão podem ser medidos não apenas em termos de acurácia, custos de classificação e precificação de spreads, mas também em termos de maximização do valor econômico das empresas alavancadas. Empresas que são avaliadas por modelos mais acurados podem adotar estruturas de capital mais eficientes, ajustando sua alavancagem ao perfil de risco, o que gera mais valor para acionistas e credores.

A fórmula simplificada do V^L é:

$$V^L = V^U + PV(Deduções\ tributárias\ das\ despesas\ com\ juros) - PV(\ Custos\ de\ dificuldades\ financeiras),$$

Onde;

- V^U: Valor da empresa sem dívida;
- $PV(Deduções\ tributárias\ das\ despesas\ com\ juros)$: Valor presente das economias fiscais geradas pelos benefícios fiscais da dívida;
- $PV(\ Custos\ de\ dificuldades\ financeiras)$: Valor presente dos custos diretos e indiretos de falência, ponderados pela probabilidade de inadimplência $P(Y = 1|S = t)$.

Modelos de previsão mais eficientes impactam diretamente o (i) benefício fiscal da dívida ao identificar empresas de menor risco, possibilitando maior uso de dívida, maximizando o benefício fiscal e (ii) custos esperados de falência, pois refletem com mais precisão

$P(Y = 1|S = t)$, minimizando os custos de falência e ajustando o perfil de risco da estrutura de capital.

Assim, o impacto dos modelos na previsão de falência pode ser avaliado em termos do V^L, traduzindo ganhos de precisão em valor econômico tangível.

Esse estudo utilizou as quatro abordagens para avaliar os modelos de previsão de falência: AUC, que mede a capacidade discriminatória entre empresas solventes e insolventes; CECI, que considera a assimetria dos custos de classificação incorreta; Spread de Crédito, que traduz as previsões em ajustes práticos de precificação de risco; e Valor da Empresa Alavancada (V^L), que avalia o impacto estratégico das previsões na estrutura de capital e no valor econômico das empresas. Essa combinação permite uma análise abrangente que integra aspectos estatísticos e econômicos, alinhados à relevância prática dos modelos para a economia.

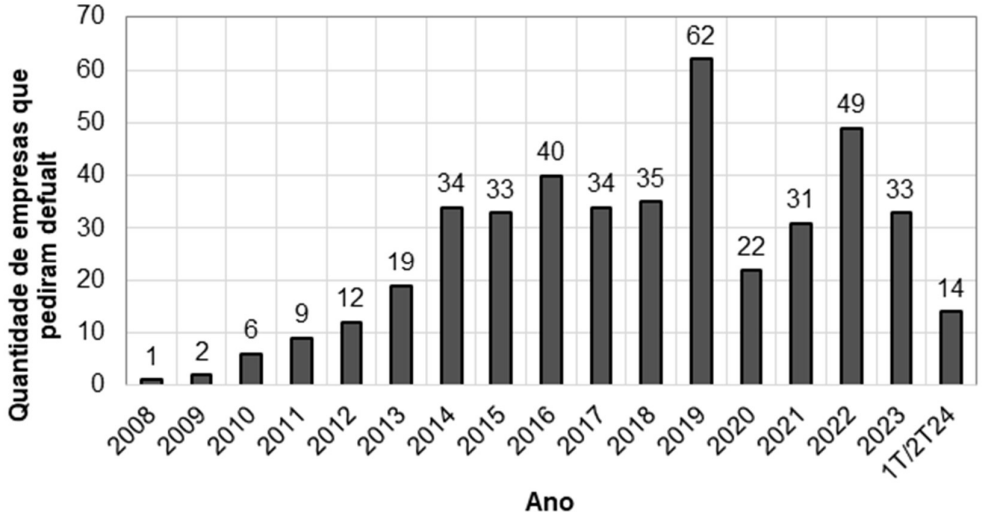
3 DADOS E METODOLOGIA

3.1 Dados e definição de default

Nesta dissertação, foram coletados dados trimestrais de empresas brasileiras, de capital aberto e de capital fechado, utilizando a plataforma Valor PRO (Valor Pro, 2024) que é um serviço de informações da Editora Globo. A amostra inclui empresas com dados financeiros disponíveis entre o segundo trimestre de 2008 (2T2008) e o segundo trimestre de 2024 (2T2024). A base de dados contém um total de 7.180 empresas, das quais 436 entraram em default em algum trimestre dentro do período analisado.

A definição de default ou dificuldades financeiras foi estabelecida com base nos critérios de classificação da Receita Federal do Brasil (RFB), conforme registrado na Rede Nacional para a Simplificação do Registro e da Legalização de Empresas e Negócios (REDESIM). Empresas com a situação cadastral especial classificadas como "Início de Recuperação Judicial" (situação especial com código 418), "Decretação de Falência" (código 405) e "Início de Liquidação Extrajudicial" (código 417) foram consideradas empresas que entraram em default. As datas de default de 325 empresas foram obtidas por meio de uma API privada CNPjá (2024) conectada ao site da REDESIM. As datas dos demais eventos que completam 436 defaults foram obtidas diretamente do site da RFB (2024), ou em colaboração com os autores Melo, Bortoluzzo et. al (2022) ou com a Associação brasileira de jurimetria ABJ (2017). As empresas classificadas na situação cadastral como "Ativa" (situação cadastral código 2) e sem registro de informação em situação cadastral especial, foram consideradas empresas que não entraram em default ou não tiveram dificuldades financeiras no período analisado.

Figura 4: Evolução do número de empresas com evento default entre 1T08 e 2T24.

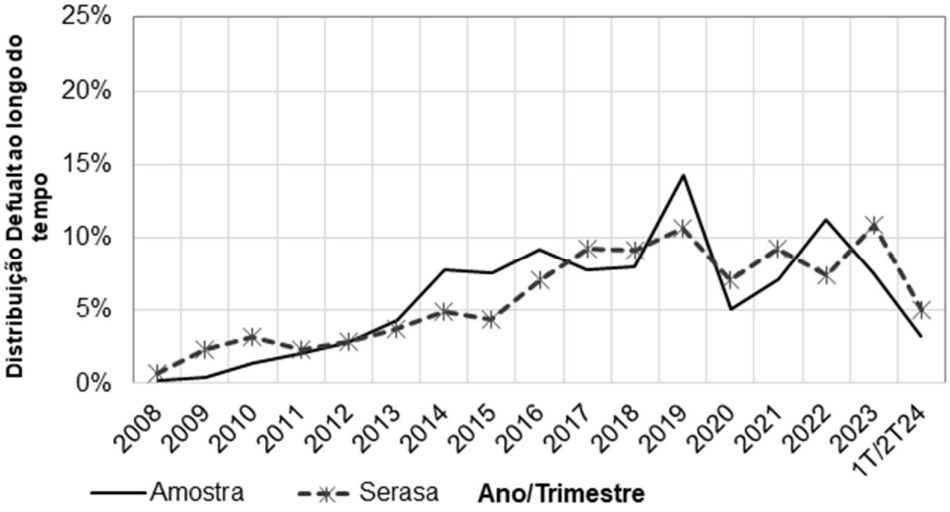


Fonte: AUTOR (2024)

Observa-se na Figura 4 a evolução do número de empresas que entraram em default na amostra entre o 2T2008 e 2T2024, destacando um aumento significativo de defaults em 2019 e em 2022. Em 2019 o aumento pode ser atribuído principalmente ao pedido de recuperação judicial do grupo "Odebrecht", decorrente de sua participação no esquema de corrupção popularmente conhecido como "Operação Lava Jato" da Polícia Federal brasileira. Empresas pertencentes a esse grupo empresarial representaram cinco defaults nesse período. Além disso, no mesmo ano, outras 13 empresas do setor de construção civil, que atuam no mesmo ramo da Odebrecht, também entraram com pedido de recuperação judicial. Em 2022 outro esquema de corrupção desencadeou em um só trimestre o pedido de recuperação judicial de 16 empresas do conglomerado empresarial "Grupo João Santos". Esses pedidos de recuperação judicial (RJ) foram decorrentes também de uma investigação da Polícia Federal conhecida como "Operação Background" amplamente divulgado pela mídia nacional.

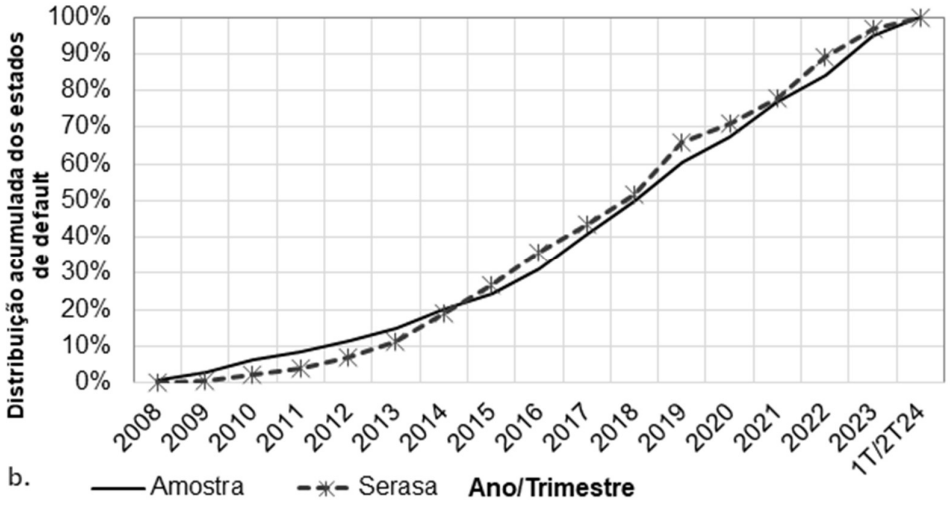
Foi realizada uma análise comparativa da amostra utilizada neste estudo com dados de uma base externa, extraída do site da Serasa (2024), uma das maiores *Datatechs* do Brasil, que coleta informações sobre empresas em dificuldades financeiras desde 1991.

Figura 5: Distribuição dos pedidos de default dos dados da amostra e dados Serasa.



Fonte: (AUTOR, 2024).

Figura 6: Distribuição acumulada de default das empresas da amostra.



Comparativo do percentual acumulado de todas as observações default entre a amostra desse estudo e amostra do SERASA. Fonte: Serasa e dados da amostra desse estudo. Fonte: (AUTOR, 2024).

Na Figura 5 observa-se a comparação, apresentando duas curvas: a linha preta e contínua representa a distribuição das empresas que entraram em default na amostra desse estudo, enquanto a linha tracejada mostra a distribuição de 6.645 empresas que tiveram pedidos de RJ deferidos pela justiça na base de dados da Serasa. Observa-se que as duas curvas exibem uma distribuição muito similar, sugerindo que a amostra utilizada neste estudo é semelhante aos dados de default coletados pela Serasa. Na Figura 6 observa-se a distribuição acumulada ao longo do tempo dos eventos de default e fica evidenciado a diferença de distribuição a partir de 2019 e 2022. Apesar dessas diferenças pontuais a

similaridade entre as duas curvas (amostra e Serasa) reforça a representatividade dos dados deste estudo e a sua capacidade de refletir o comportamento do mercado brasileiro de insolvência empresarial.

Além da validação temporal da representatividade da amostra, foi realizada uma caracterização detalhada das empresas incluídas neste estudo, com base em sua natureza jurídica e porte econômico. Os dados, apresentados nas Tabelas 12 e 13 no Apêndice C, indicam uma predominância de empresas de médio e grande porte, que correspondem a mais de 94% da amostra total, com baixa participação relativa de micro e pequenas empresas. Quanto à natureza jurídica, observa-se concentração em empresas privadas com fins lucrativos, especialmente sociedades limitadas e anônimas. Essa composição decorre, em grande medida, da cobertura da base Valor PRO, que tende a incluir empresas com maior formalização e disponibilidade de informações contábeis. Embora tal perfil assegure consistência e qualidade aos dados utilizados, também implica uma limitação metodológica relevante, na medida em que a amostra não representa adequadamente o universo das micro e pequenas empresas, segmento que, conforme demonstrado pelos dados recentes da Serasa Experian, apresentou o maior crescimento proporcional nos pedidos de recuperação judicial no país (89,6% no período analisado).

Entretanto, é importante destacar que o fenômeno da elevação dos pedidos de recuperação judicial não se restringe às micro e pequenas empresas. As médias e grandes empresas também apresentaram aumentos substanciais, com destaque para as grandes, cujo crescimento ultrapassou 50% entre o primeiro semestre de 2023 e o mesmo período de 2024. Esse contexto reforça a relevância de modelos preditivos que sejam robustos e adaptáveis a diferentes perfis empresariais, bem como a necessidade de considerar as limitações e possibilidades decorrentes da composição da base de dados utilizada.

3.2 Variáveis explicativas

As variáveis explicativas foram selecionadas com base em estudos de Sigrist e Leuenberger (2022), abreviado como “L.S 2023” além das variáveis propostas no trabalho seminal de Altman (1968) abreviado como “Altman Z1 e no estudo complementar de Altman (1984), abreviado “Altman Z2”, que incluiu a análise de dados brasileiros. A seleção resultou em um conjunto de 12 índices financeiros, 1 índice de tamanho relativo e 4 *dummies* setoriais, que estão descritos na Tabela 1. Essas variáveis

podem ser analisadas em categorias essenciais para a análise econômica e financeira das empresas em 4 dimensões, (i) lucratividade (LL_AT, RL_AT e EB_AT), que medem a capacidade de gerar e reter lucros; (ii) liquidez (CX_AT, AC_PC e CG_AT), que avaliam a capacidade de cumprir obrigações de curto prazo; (iii) estrutura de capital e alavancagem (PT_AT, PL_AT e DB_PT) e (iv) resultado operacional (RO_AT, FO_PT e VD_AT).

Tabela 3: Variáveis preditores

Definição	Código da variável explicativa utilizada nesse trabalho.	Variável utilizada no modelo Altman Z1	Variável utilizada no modelo Altman Z2	Variável utilizada no trabalho de L.S - 2023
Lucro Líquido/Ativos Totais	LL_AT	não utilizado	não utilizado	X1
Vendas/Ativos Totais	VD_AT	X1	X1	X2
EBIT/Ativos Totais	EB_AT	X2	X2	X3
Fluxo de caixa operacional /Passivos totais.	FO_PT	não utilizado	não utilizado	X4
Resultado Operacional/Ativos Totais	RO_AT	não utilizado	não utilizado	não utilizado
Tamanho relativo (tamanho de cada empresa em relação ao setor a que pertence por trimestre)	TR	não utilizado	não utilizado	X5
Passivos Totais/Ativos Totais	PT_AT	não utilizado	não utilizado	X6
Capital de Giro/Ativos Totais	CG_AT	não utilizado	X4	X7
Reservas de Lucro/Ativos Totais	RL_AT	X3	não utilizado	X8
Ativos Circulante/Passivos Circulante	AC_PC	não utilizado	não utilizado	X9
Caixa/Ativos Totais	CX_AT	não utilizado	não utilizado	X10
Patrimonio Líquido /Ativos Totais	PL AT	X5	X5	X11
Dívida Bruto/Passivos Totais	DB_PT	X6	não utilizado	não utilizado
Dummies Setoriais	manu_mine	não utilizado	não utilizado	X12
	trans_util	não utilizado	não utilizado	X13
	constr	não utilizado	não utilizado	não utilizado
	diversos	não utilizado	não utilizado	X15

Comparação das variáveis utilizada nesse estudo com as variáveis que foram utilizadas nos trabalhos seminais de Altman (1968, 1984) e no trabalho de Sigrist e Leuenberger (2022). Fonte: AUTOR (2024).

3.3 Modelo teórico e econométrico

Para estruturar um modelo teórico adequado à previsão de falências corporativas essa dissertação utiliza como referencial metodológico o trabalho de Sigrist e Leuenberger (2022). O objetivo é modelar como fatores como lucratividade, liquidez e alavancagem influenciam a probabilidade de uma empresa entrar em default em um horizonte de tempo específico. O modelo teórico usado pode ser expresso da seguinte forma:

$$P(\tau_i = t|X_{ijt}) = F\left(X_{ijt}\right)$$

Onde:

- X_{ijt}: é o valor observado da empresa i, da variável explicativa j, no tempo t.
- $P(\tau_i = t|X_{ijt})$: a probabilidade condicional de default τ da empresa *i* no tempo *t.*
- $F(.)$: é uma função de ligação que pode ser uma função logística (no caso de uma regressão logística) ou uma função de previsão mais complexa usada em modelos de aprendizado de máquina;

O objetivo é estimar as probabilidades de default em cada período e analisar como diferentes variáveis financeiras impactam o risco de falência de uma empresa. A metodologia é baseada no conceito da probabilidade condicional de default *forward*. A variável $\tau_i \in \mathbb{N}^+$representa o instante de tempo de default da empresa i.

Assume-se que as variáveis explicativas e o evento de default são observados em intervalos de tempos discretos $t \in \mathbb{N}, 0 \leq t_{0i} \leq t \leq T_i \in \mathbb{N}^+$, onde $t_{0i} \in \mathbb{N}$ representa o momento em que a empresa i entra no conjunto de empresas observáveis e T_i é o último tempo de observação da empresa i, que pode ser o tempo de default τ_i, uma saída por indisponibilidade de dados, fusão ou aquisição, ou o final do período de análise.

É importante explicitar t_{0i}, uma vez que nem todas as empresas estão disponíveis para observação durante todo o intervalo de tempo analisado. Com base nessas variáveis, calcula-se a probabilidade condicional de default em cada intervalo, levando em consideração se a empresa já entrou em default ou não até o período t.

A probabilidade de default condicional *forward* é definida como a probabilidade de uma empresa entrar em default k períodos após o t, onde $k \leq j$ e $j \in \mathbb{N}^+$determina o horizonte de previsão. Em outras palavras, a probabilidade condicional *forward* mede o risco de uma empresa falir dentro do intervalo de tempo $(t + k - 1, t + k]$ dado que ela não entrou em default até o tempo $t + k - 1$. Formalmente, essa probabilidade condicional *forward* é representada da seguinte forma:

$$P_t(\tau_i = t + k|\tau_i > t + k - 1).$$

A probabilidade *forward* calcula o risco de uma empresa falir em um período específico futuro, dado que ela tenha sobrevivido até o momento anterior.

O objetivo é modelar as probabilidades de default a partir de funções de predição que relacionam as variáveis explicativas com a chance de falência. Para isso, utilizou-se a função $G_f^k(X_t^i)$, que é a função de predição que mapeia as variáveis explicativas da

empresa i no tempo t para as probabilidades *forward*, ou seja, é a função que recebe como entrada as variáveis explicativas X_t^i da empresa i no tempo t e devolve a probabilidade de default entre [0, 1] para o próximo período. Logo, $G_f^k(X_t^i)$ é a função que modela a probabilidade condicional *forward*, ou seja, a probabilidade de a empresa i entrar em default no intervalo de tempo $(t+k-1, t+k]$, dado que ela não faliu até o tempo $t+k-1$. Formalmente, essa probabilidade é expressa como:

$$G_f^k(X_t^i) = P_t(\tau_i = t+k | \tau_i > t+k-1)$$

Para cada empresa i definiu-se $\mathcal{F}$ como o conjunto de informações observáveis. O subconjunto $\mathcal{F}_t \subset \mathcal{F}$ representa uma seleção de informações em tempo discreto e é definido pela função:

$$\mathcal{F}_t = \sigma\left(X_i^t, 1_{\{\tau_i < t\}};\ 1 \leq i \leq N, t_{0i} \leq t \leq T_i\right)$$

onde:

- $\sigma(.)$ representa a álgebra gerada pelas variáveis indicadas, ou seja, o menor conjunto de eventos que contém todas as informações disponíveis até o tempo t, capturando tanto as variáveis explicativas quanto os eventos de default observados;
- X_i^t representa o conjunto de todas as variáveis explicativas observáveis para a empresa i no tempo t;
- $1_{\{\tau_i < t\}}$ é uma função indicadora que assume o valor 1 se a empresa i entrou em default até o tempo t, e 0 caso contrário;
- A condição $1 \leq i \leq N$ indica que a seleção $\mathcal{F}_t$ considera todas as empresas i dentro do conjunto de N empresas observadas;
- A condição $t_{0i} \leq t \leq T_i$ define o intervalo de tempo no qual a empresa i está sendo observada, começando no tempo t_{0i}, quando a empresa entra no conjunto de observação, e terminando no tempo T_i. Esse tempo T_i pode ser o momento em que a empresa entra em default, sai do conjunto de observação por outro motivo, ou o fim do período total de análise.

Na Tabela 4 observa-se como os dados devem ser organizados para o cálculo das probabilidades de default *forward*, com base nos conceitos apresentados na metodologia. O exemplo considera uma empresa fictícia (Empresa A), observada ao longo de 10 períodos e que entrou em default no período 9.

Tabela 4: Estrutura de dados para cálculos de probabilidades condicionais *foward*.

Variáveis observadas					Informação ilustrativas, para demonstração do método						Variáveis respostas $Y_t^{(k)}$		
t	i	τ_i	X_{1i}	X_{2i}	t+k-1 (k=1)	t+k-1 (k=2)	t+k-1 (k=3)	t+k (k=1)	t+k (k=2)	t+k (k=3)	k=1	k=2	k=3
1	A	0	0,01	0,01	1	2	3	2	3	4	0	0	0
2	A	0	0,04	-0,02	2	3	4	3	4	5	0	0	0
3	A	0	0,02	-0,05	3	4	5	4	5	6	0	0	0
4	A	0	0,02	0,02	4	5	6	5	6	7	0	0	0
5	A	0	0,05	0,08	5	6	7	6	7	8	0	0	0
6	A	0	0,07	0,07	6	7	8	7	8	**9**	0	0	**1**
7	A	0	-0,02	-0,06	7	8	**9**	8	**9**	10	0	**1**	NA
8	A	0	-0,06	0,02	8	**9**	10	**9**	10	11	**1**	NA	NA
9	A	**1**	0,01	0,04	**9**	10	11	10	11	12	NA	NA	NA
10	A	0	0,05	-0,06	10	11	12	11	12	13	NA	NA	NA

- Período (t): Representa o tempo discreto em que a empresa foi observada. Em cada período, são coletadas informações sobre as variáveis explicativas e o estado da empresa em relação ao default (se já ocorreu ou não);
- τ_i (Indicador de Default): Indica se a empresa entrou em default em um período específico, onde o valor 1 indica que a empresa entrou em default naquele período, e 0 indica que permaneceu solvente;
- Variáveis Explicativas (X_{1i}, X_{2i},..., X_{ni}): são as variáveis financeiras observáveis no tempo t;
- Período ($colunas\ t+k-1$ e $t+k$): referem-se aos períodos anteriores e posteriores ao período atual, sendo usadas para calcular indicar o estado de default em cada horizonte;
- Colunas Y ($k=1$, $k=2, k=3$): representam as variáveis de resposta para os modelos de probabilidade *forward*. Cada coluna indica se a empresa entrou em default exatamente em k períodos à frente, com base nas condições anteriores. Para cada k , a coluna correspondente terá valor 1 se a empresa entrou em default exatamente k períodos após o tempo t, e 0 caso contrário, ou seja,

  $Y_t^{(k)} = \begin{cases} 1, & se\ \tau_i = t+k \\ 0, & se\ \tau_i > t+k \\ NA, & se\ \tau_i \leq t+k-1 \end{cases}$, onde $Y_t^{(k)}$representa a variável de interesse no tempo t no horizonte k;
- Preenchimento após o default: Uma vez que a empresa entra em default em um determinado período t, todas as observações subsequentes são desconsideradas (marcadas como "NA"). Isso reflete a hipótese de que, após o default, a empresa

não pode "reentrar" em default, e, portanto, suas observações deixam de ser relevantes para o modelo a partir desse ponto.

A partir do modelo da Tabela 4, os diferentes métodos de classificação (Altman z-score, Random Forest e outros métodos) foram ajustados utilizando as variáveis explicativas X_{ni}. As colunas binárias intituladas como $k = 1$, $k = 2, k = 3$ indicam se a empresa entrou em default em diferentes horizontes *forward* e representam as variáveis de interesse. Para cada horizonte, foi ajustado um modelo e as probabilidades *forward* de default foram calculadas com base nesses resultados.

Na Seção 3.6, os modelos probabilísticos de classificação para a determinação das funções $G_f^k(X_t^i)$ são explorados em maior detalhe. Os modelos discutidos são Altman Z-Score de 2005 para mercados emergentes Altman EMS (2005), baseado em análise discriminante multivariada, que continua a ser uma referência clássica para a previsão de falências e os modelos de aprendizado de máquina; Regressão logística; Random Forest, Light Gradient Boosting e CatBoost.

3.4 Estatísticas descritivas

Na Tabela 5 apresentam-se as estatísticas descritivas comparativas entre as 6.744 empresas que não entraram em default e as 436 que apresentaram default ao longo do período analisado. São exibidas as médias das variáveis preditoras para cada grupo, acompanhadas dos valores-p obtidos por meio do teste t de Welch, que avaliam a significância estatística das diferenças observadas. As variáveis de lucratividade, liquidez e alavancagem mostram diferenças relevantes entre os dois grupos.

De modo geral, as empresas que não apresentaram default exibiram melhores indicadores de lucratividade, com médias positivas tanto para o lucro líquido sobre ativos totais (LL_AT) quanto para o EBIT sobre ativos (EB_AT), enquanto as empresas em default registraram prejuízos. A capacidade de reter lucros (RL_AT) também foi consideravelmente maior nas empresas solventes, evidenciando padrões financeiros diferenciados entre os grupos.

Os resultados do teste t de Welch indicam que, para praticamente todas as variáveis analisadas, existem diferenças estatisticamente significativas entre as médias dos dois grupos, com valores-p inferiores a 0,05. Esse resultado reforça a capacidade preditiva dos

indicadores financeiros selecionados e sua importância na discriminação entre empresas solventes e insolventes.

Tabela 5 Estatísticas comparativas entre empresas solventes e em default com teste t de significância

Variável	Média Solventes	Média Default	p-valor (teste t)
Lucratividade			
LL_AT	0,0066	(0,0196)	p < 1e-10
EB_AT	0,0163	(0,0019)	p < 1e-10
RL_AT	0,1631	0,0777	p < 1e-10
Liquidez			
AC_PC	3,9781	2,7656	p = 0,0459
CX_AT	0,0911	0,0460	p < 1e-10
CG_AT	0,1266	(0,0880)	p < 1e-10
Alavancagem			
PT_AT	0,5939	1,0042	p < 1e-10
PL_AT	0,4060	(0,0044)	p < 1e-10
DB_PT	0,1100	0,2404	p < 1e-10
Resultado Operacional/Eficiência de vendas			
VD_AT	0,3061	0,2792	p = 0,0434
FO_PT	0,0205	0,0014	p < 1e-10
RO_AT	0,0106	(0,0183)	p < 1e-10

Dentre os destaques, as variáveis de lucratividade — como LL_AT e EB_AT — apresentaram médias significativamente mais elevadas nas empresas solventes, indicando sua relevância como sinalizadores precoces de deterioração financeira. O mesmo padrão se verifica em variáveis de liquidez, como o caixa sobre ativos totais (CX_AT), e em indicadores de estrutura de capital, como o passivo total sobre ativos (PT_AT) e o patrimônio líquido sobre ativos (PL_AT), cujas diferenças entre os grupos refletem a fragilidade patrimonial das empresas que faliram.

Por outro lado, o único indicador cuja diferença apresentou significância estatística marginal foi o AC_PC (Ativo Circulante / Passivo Circulante), cuja elevada variabilidade entre as empresas analisadas pode ter atenuado sua capacidade discriminatória isolada.

Esses resultados fortalecem a validade empírica da base, demonstrando que as variáveis financeiras analisadas são efetivas em capturar diferenças estruturais relevantes entre os

grupos. Para uma compreensão ainda mais detalhada dos dados — incluindo desvios padrão, quartis e valores extremos.

Tabela 6: Estatísticas das variáveis preditoras das empresas solventes.

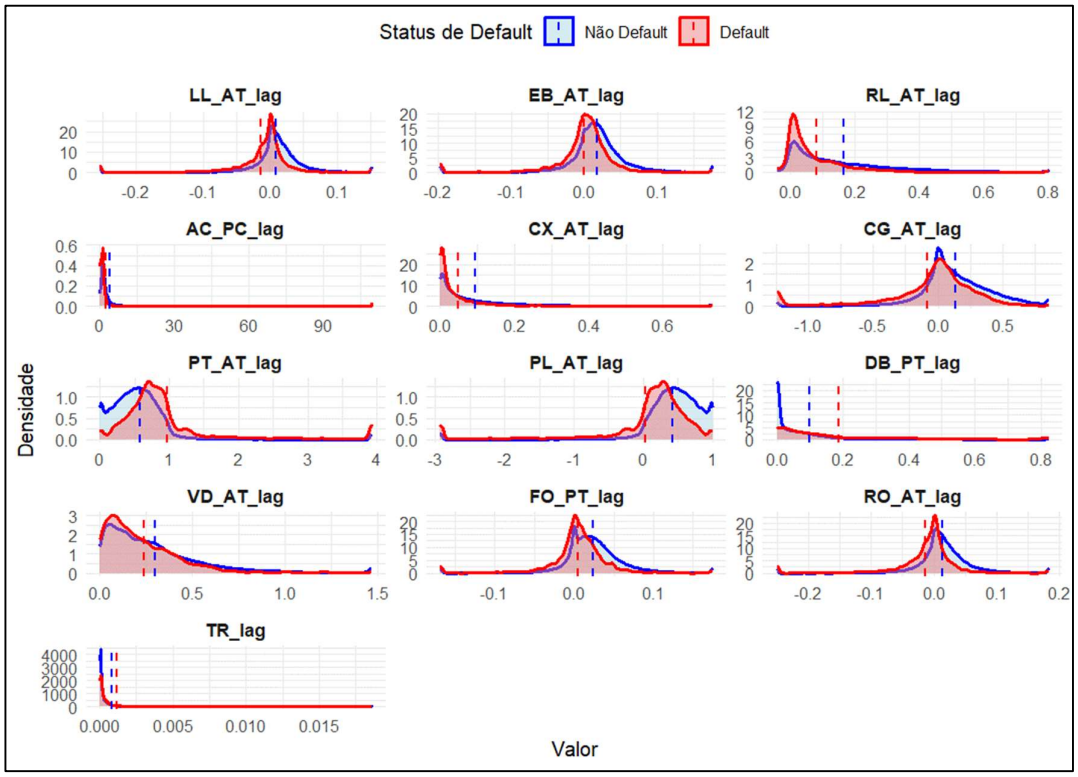
Variável	Média	Desv.Pad	Mínimo	Q25	Mediana	Q75	Máximo
Lucratividade							
LL_AT	0,0066	0,0498	(0,2928)	(0,0043)	0,0079	0,0245	0,1674
EB_AT	0,0163	0,0466	(0,2062)	(0,0005)	0,0153	0,0340	0,1823
RL_AT	0,1631	0,1812	(0,0667)	0,0252	0,1014	0,2429	0,8365
Liquidez							
AC_PC	3,9781	13,0625	0,0153	0,9266	1,4514	2,5061	122,3919
CX_AT	0,0911	0,1339	0,0000	0,0095	0,0381	0,1141	0,7339
CG_AT	0,1266	0,3106	(1,4613)	(0,0124)	0,1040	0,2963	0,8716
Alavancagem							
PT_AT	0,5939	0,5313	0,0001	0,3085	0,5376	0,7534	4,3988
PL_AT	0,4060	0,5315	(3,4033)	0,2466	0,4624	0,6915	0,9999
DB_PT	0,1100	0,1743	0,0000	0,0000	0,0368	0,1308	0,9680
Resultado Operacional/Eficiência de vendas							
VD_AT	0,3061	0,2822	0,0004	0,1025	0,2344	0,4155	1,4675
FO_PT	0,0205	0,0454	(0,2091)	0,0009	0,0185	0,0388	0,1838
RO_AT	0,0106	0,0537	(0,2848)	(0,0051)	0,0099	0,0309	0,1958

Tabela 7: Estatísticas das variáveis preditoras das empresas que entraram em default.

Variável	Média	Desv.Pad	Mínimo	Q25	Mediana	Q75	Máximo
Lucratividade							
LL_AT	(0,0196)	0,0644	(0,2928)	(0,0266)	(0,0044)	0,0051	0,1674
EB_AT	(0,0019)	0,0506	(0,2062)	(0,0123)	0,0021	0,0176	0,1823
RL_AT	0,0777	0,1013	(0,0667)	0,0111	0,0477	0,1139	0,8365
Liquidez							
AC_PC	2,7656	12,2050	0,0153	0,5142	1,0854	1,5320	122,3919
CX_AT	0,0460	0,0827	0,0000	0,0040	0,0167	0,0523	0,7339
CG_AT	(0,0880)	0,4411	(1,4613)	(0,1961)	0,0210	0,1423	0,8716
Alavancagem							
PT_AT	1,0042	0,8623	0,0001	0,5964	0,8055	0,9866	4,3988
PL_AT	(0,0044)	0,8629	(3,4033)	0,0134	0,1945	0,4036	0,9999
DB_PT	0,2404	0,2572	0,0000	0,0424	0,1425	0,3785	0,9680
Resultado Operacional/Eficiência de vendas							
VD_AT	0,2792	0,2679	0,0004	0,0992	0,1959	0,3809	1,4675
FO_PT	0,0014	0,0493	(0,2091)	(0,0090)	0,0048	0,0211	0,1838
RO_AT	(0,0183)	0,0655	(0,2848)	(0,0286)	(0,0052)	0,0068	0,1958

Observa-se na Figura 7 a distribuição das variáveis entre os dois grupos. Fica evidente que as empresas que não entraram em default apresentam distribuições mais deslocadas para a direita nas variáveis de lucratividade, como LL_AT e EB_AT, indicando maior capacidade de geração de lucros. Entre as variáveis de liquidez, a distribuição de CG_AT mostra uma clara diferenciação: as empresas em default estão mais concentradas em valores negativos, indicando dificuldades na conversão de ativos em capital de giro. A alavancagem também diferencia os grupos, com empresas em default apresentando maior dependência de capital de terceiros, o que pode aumentar o risco em cenários de altos custos de financiamento. A geração de fluxo de caixa operacional (FO_PT) é mais alta nas empresas solventes.

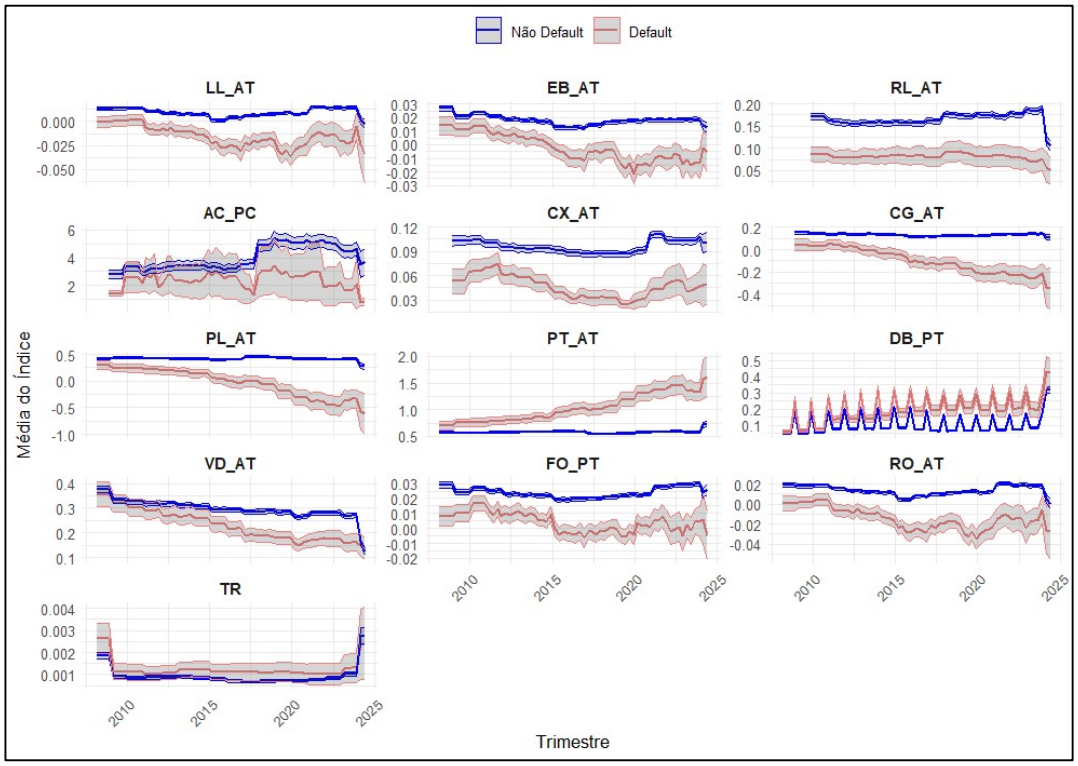
Figura 7: Distribuição das variáveis preditoras.



Cada gráfico da figura tem as escalas dos eixos definidas pelo critério da própria variável. As linhas tracejadas em vermelho representam a média das empresas que entraram em default e azul a média das empresas solventes. Fonte: (AUTOR, 2024).

Na Figura 8 observa-se a evolução das médias de cada variável ao longo do período analisado, comparando as empresas que entraram em default com aquelas que não entraram. Em todas as variáveis, as empresas que faliram apresentam indicadores inferiores ao longo do tempo, demonstrando que seu desempenho financeiro foi consistentemente mais fraco.

Figura 8: Médias das variáveis preditoras ao longo do tempo.



Índices financeiros de empresas solventes (linhas azuis) e em default (linhas vermelhas) ao longo dos trimestres. As bandas de preenchimento ao redor das linhas representam os intervalos de confiança de 95%, calculados com base no erro padrão e no tamanho da amostra em cada grupo. Os gráficos são organizados por índice financeiro, com escalas ajustadas para cada variável, destacando a incerteza das médias e as diferenças entre os grupos. Fonte: (AUTOR, 2024).

A análise da matriz de correlação que podem ser observadas na Figura 9 revela que a lucratividade (LL_AT) está fortemente correlacionada com outras medidas de desempenho, como RO_AT (0,96) e FO_PT (0,73), sugerindo que empresas com maior lucro líquido também apresentam melhor desempenho operacional. Por outro lado, PT_AT (passivo total sobre ativos totais) tem uma correlação negativa significativa com LL_AT (-0,41), evidenciando o impacto negativo da alavancagem sobre a lucratividade.

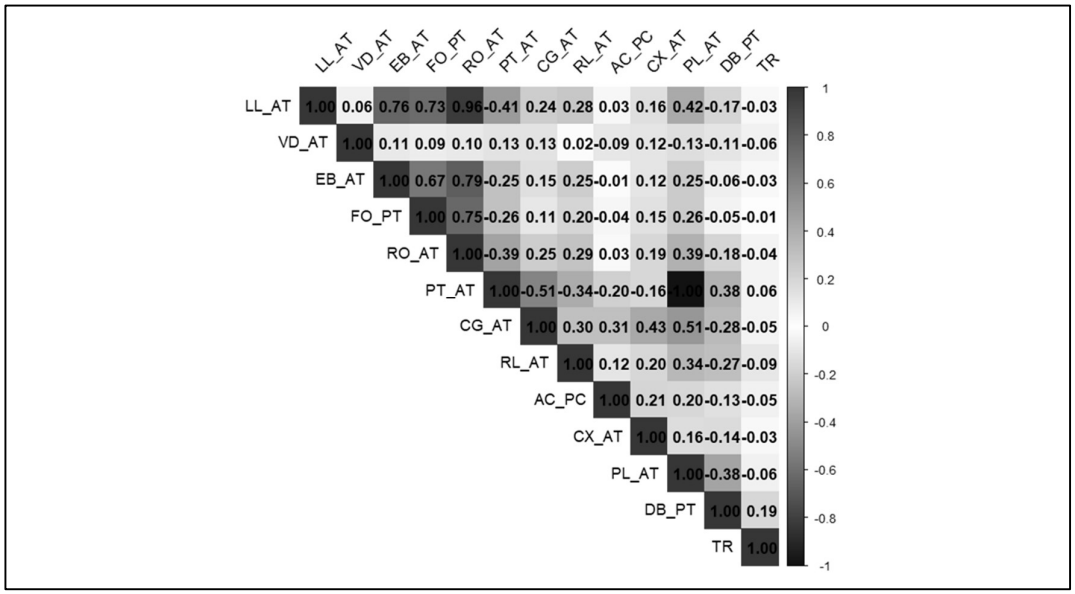
Quanto às variáveis de liquidez, CX_AT (caixa sobre ativos totais) apresenta correlações moderadas com o desempenho financeiro, como LL_AT (0,16) e RO_AT (0,19), indicando que empresas com maior caixa tendem a ter melhores resultados operacionais. Já a alavancagem, medida por DB_PT (dívida bruta sobre patrimônio), tem uma correlação negativa com variáveis de lucratividade, como RL_AT (-0,27), sugerindo que empresas mais endividadas têm maior dificuldade em reter lucros. No geral, essas correlações mostram que a lucratividade e o desempenho operacional estão positivamente

associados, enquanto a alavancagem exerce um impacto negativo sobre a saúde financeira das empresas.

Além disso, a correlação extremamente forte entre PT_AT e PL_AT (-1,0) indica uma redundância entre essas variáveis, uma vez que ambas refletem aspectos semelhantes da estrutura de capital. As correlações fortes entre LL_AT e RO_AT (0,96), assim como entre EB_AT e LL_AT (0,76), serão cuidadosamente tratadas nos modelos preditivos, uma vez que essas variáveis estão diretamente relacionadas à lucratividade e ao desempenho operacional.

Para o modelo de regressão logística, a avaliação estatística detalhada das correlações entre as variáveis foi desconsiderada, dado que o foco principal está na precisão preditiva da classificação, e não na explicação causal. Assim, a seleção e o tratamento das variáveis priorizaram a eficácia do modelo em classificar corretamente os eventos de falência

Figura 9: Matriz de correlação das variáveis preditoras.



Fonte: (AUTOR, 2024).

3.5 Divisão dos dados em Treino, Validação e Teste

A divisão dos dados foi realizada em três conjuntos; (i) treino, (ii) validação e (iii) teste. O conjunto de treino abrange os trimestres de 2T2008 até 1T2016, contendo 116.725 observações de empresas com dados não faltantes, representando 43% do total de observações. O conjunto de validação abrange o período de 2T2016 até 3T2019, com

71.722 observações, equivalendo a 27% do total de observações. O conjunto de teste cobre o período de 4T2019 até 2T2024, contendo 80.238 observações, o que corresponde a 30% do total de observações. No total, a amostra completa contém 268.685 observações de empresas ao longo de todos os períodos. Essa divisão assegura que 70% das observações estejam no conjunto de treino e validação, dos quais 62% no conjunto de treino, 38% para validação e sobre a amostra total 30% no conjunto de teste.

Com essa estrutura, os modelos foram ajustados nos dados de treino, avaliados nos dados de validação, e posteriormente ajustados nos dados de teste.

Para lidar com possíveis dados faltantes entre períodos, foi implementada a imputação dos valores ausentes utilizando o último valor disponível. Nos casos em que os dados trimestrais não estavam disponíveis, mas os dados anuais sim, os dados de resultado como Lucro líquido, ou EBIT foram trimestralizados e os dados de balanço como Ativo Circulante foram considerados iguais ao do último trimestre do ano reportado. Após esse processo de imputação, os dados foram *winsorizados* nos percentis 1% e 99% para mitigar o impacto de outliers, conforme as práticas descritas por Sigrist e Leuenberger (2022) e Tian et al. (2015).

Como os relatórios financeiros são, em geral, divulgados com atraso em relação ao período de referência, aplicou-se uma defasagem (*lag*) de 1 trimestre em todos os índices financeiros. Dessa forma, os dados contábeis utilizados no modelo referem-se ao trimestre anterior ao evento observado.

3.6 Métodos de classificação

Nesta seção, são apresentados os principais aspectos dos modelos de classificação utilizados para prever o risco de falência das empresas. Foram aplicados os modelos de altman z-score para mercados emergentes, regressão logística, random forest, ligth gradient boosting, e catboost. Cada um dos modelos propostos é utilizado para representar a função $G_f^k(.)$, que captura as probabilidades *forward* de default.

3.6.1 Altman Z-Score para mercados emergentes (EMS)

O modelo Z-Score para mercados emergentes (EMS), desenvolvido por Altman (2005), é um dos modelos clássicos utilizados para previsão de falências. Baseado em uma análise

discriminante multivariada, ele combina diversos indicadores financeiros, como liquidez, lucratividade, e alavancagem, para calcular um score que prevê a probabilidade de uma empresa falir.

A fórmula do Z-Score é dada por:

$$Z = 3{,}25 + 6{,}56 \times \left(\frac{Capital\ de\ Giro}{Ativo\ Total}\right) + 3{,}26 \times \left(\frac{Reservas\ de\ Lucro}{Ativo\ Total}\right) + 6{,}72 \times \left(\frac{EBIT}{Ativo\ Total}\right) + 1{,}05 \times \left(\frac{Patrimônio\ Líquido}{Ativo\ Total}\right)$$

O Z-Score para mercados emergentes (EMS) é uma métrica consolidada e de fácil interpretação, que classifica as empresas em diferentes zonas de risco (zona de falência, zona de alerta e zona de segurança) com base no valor calculado. Se Z-Score for inferior a 1,10, a empresa é classificada como em dificuldades financeiras, entre 1,10 e 2,60, a empresa está em uma zona de alerta, superior a 2,60, a empresa é considerada financeiramente segura, com baixo risco de falência.

Nesse estudo, o z-score do modelo de Altman foi transformado em uma probabilidade contínua entre 0 e 1 utilizando uma regressão logística ajustada para prever a probabilidade de default. Essa abordagem permite comparar diretamente as probabilidades previstas pelo modelo de Altman com outros modelos preditivos. A transformação é realizada ajustando a regressão logística na forma:

$$P(\tau_i = t|X_{it}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 . z_score)}}$$

Onde β_0e $\beta_1$são os coeficientes estimados pelo modelo e $P(\tau_i = t|X_{it})$ é a probabilidade condicional de falência da empresa i no tempo t. Isso assegura que o z-score seja convertido em uma métrica contínua de probabilidade, mais adequada para avaliação comparativa e análise de desempenho preditivo.

3.6.2 Regressão Logística

A Regressão Logística Cox (1958) é um método amplamente utilizados na previsão de falências corporativas. Sua simplicidade e interpretação facilitada tornam-no um ponto de partida importante para modelos de classificação binária, onde o objetivo é prever se uma empresa irá falir ou não (variável dependente binária). A fórmula geral é dada por;

$$P(\tau_i = t|X_{it}) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_{1ti} + \cdots + \beta_n X_{nti})}}$$

Onde;

- $P(\tau_i = t|X_{it})$ é a probabilidade condicional de falência da empresa i no tempo t.
- X_{it} representa a variável explicativa da empresa i no tempo t.
- $\beta_0, \beta_1, \dots, \beta_n$ são os coeficientes estimados.

A regressão logística foi implementada com pacote *glmnet* (Friedman et al., 2010) da linguagem de programação R.

3.6.3 Random Forest

O Random Forest Breiman (2001) é um método de aprendizado de máquina baseado em árvores de decisão que combina a previsão de múltiplas árvores para melhorar a precisão e reduzir o risco de *overfitting*. A predição final do Random Forest é obtida pela média das predições das T árvores.

$$C(\tau_i \mid X_{it}) = \frac{1}{T}\sum_{t=1}^{T} C_t(X_{it})$$

Onde;

- $C(\tau_i \mid X_{it})$ é a classe predita (default ou não-default) da empresa i dado as variáveis observadas.
- T é o número de árvores na floresta;
- $C_t(X_{it})$ é a previsão da $t - ésima$ árvore baseada nas variáveis X_{it} da empresa i.

A fórmula indica que o Random Forest faz a previsão final ao combinar as previsões de todas as T árvores individuais. No caso de problemas de classificação, a combinação geralmente é feita por votação majoritária — a classe que aparece mais vezes entre as árvores é escolhida como a previsão final. O modelo foi implementado no software R utilizando o pacote *ranger* Wright & Ziegler, (2017).

3.6.4 Light Gradient Boosting Machine

Gradient Boosting James et al. (2023) é um método de aprendizado que constrói modelos sequencialmente, onde cada novo modelo tenta corrigir os erros cometidos pelos modelos anteriores. Em vez de ajustar diretamente os valores-alvo, o Gradient Boosting ajusta os resíduos, que representam os erros do modelo anterior. O processo começa com uma

predição inicial $F_0(X)$, e a cada iteração m um novo modelo $h_m(x)$ é ajustado para prever os resíduos. A predição final é atualizada adicionando-se esse novo modelo corrigido, ponderado por um fator de taxa de aprendizado γ, que controla a magnitude da correção em cada etapa. Dessa forma, o modelo gradualmente melhora suas previsões ao longo das iterações, reduzindo os erros progressivamente. Matematicamente, a sequência de iterações pode ser descrita da seguinte forma:

$$F_{m+1}(X_{it}) = F_m(X_{it}) + \gamma h_m(X_{it})$$

Onde:

- $F_m(X_{it})$ é a predição ajustada do risco de falência após m+1 iterações;
- $h_m(X_{it})$ é o modelo ajustado aos resíduos da iteração anterior m, baseado nos índices financeiros e nas variáveis de setor
- γ é a taxa de aprendizado, que determina o quanto o modelo ajustado deve influenciar a nova predição.

A taxa de aprendizado γ é um dos hiperparâmetros: valores menores tornam o processo mais lento, mas podem melhorar a precisão ao permitir um ajuste mais gradual. Ao final, o Gradient Boosting combina os vários modelos sequenciais em um ensemble que minimiza os erros do conjunto de dados. Para implementar o Gradient Boosting, foi utilizado pacote LigthGBM Shi et al., (2024) na linguagem R.

3.6.5 CatBoost

O CatBoost, Dorogush et al. (2018) e Prokhorenkova et al. (2017) é uma versão otimizada do Gradient Boosting, desenvolvida para lidar eficientemente com variáveis categóricas. O modelo básico do CatBoost segue a fórmula geral do Gradient Boosting, com um tratamento especial para variáveis categóricas.

$$F_{m+1}(X_{it}) = F_m(X_{it}) + \gamma h_m(X_{it}) + \sum_{categorias} codificação\ das\ categorias$$

Onde:

- $F_{m+1}(X_{it})$ é a predição atualizada para a empresa i no tempo t;
- $F_m(X_{it})$: predição anterior do modelo no passo m;
- $h_m(X_{it})$: modelo ajustado para corrigir os resíduos da iteração anterior;

- γ: a taxa de aprendizado, que controla a magnitude da correção feita em cada iteração;
- $\sum cod.categorias$: o termo que trata a codificação das variáveis categóricas, como por exemplo, transformar variáveis contínuas de índices financeiros em intervalor de valores e em seguida criar *dummies* para cada intervalo.

O CatBoost foi implementado usando o pacote catboost, Dorogush et al. (2018) e Prokhorenkova et al., (2017).

O CatBoost utiliza um método para lidar com variáveis categóricas que se diferencia do tradicional *one-hot encoding* utilizado nos outros métodos como ANN's e Gradient Boosting. No *one-hot encoding*, cada categoria de uma variável é transformada em várias colunas binárias, o que aumenta a dimensionalidade dos dados e pode causar problemas de eficiência em modelos complexos. No entanto, o CatBoost aplica uma abordagem de codificação de categorias baseada em estatísticas de *target* (variável resposta), como a média dos valores da variável resposta para cada categoria. Isso é feito de maneira incremental, utilizando permutações aleatórias dos dados, o que evita o problema conhecido como *target leakage* - quando informações futuras influenciam a previsão do modelo, distorcendo os resultados. Ao usar essa técnica, o CatBoost permite que o modelo lide com variáveis categóricas sem aumentar a complexidade computacional e sem comprometer a robustez do modelo.

No caso específico para previsão de default, as *dummies* de setor não serão transformadas em múltiplas colunas binárias através do *one-hot encoding*. Em vez disso, o CatBoost calculará de forma incremental a média dos índices financeiros (variável resposta) para cada categoria de setor, garantindo que a observação atual não influencie o cálculo dessas médias. Esse método permite que o modelo capture as interações entre os setores e os índices financeiros, sem aumentar o número de variáveis no modelo e evitando o problema de *target leakage*.

Os modelos de aprendizado de máquina apresentados nessa seção dependem de uma série de hiperparâmetros que influenciam diretamente seu desempenho e eficiência. Esses hiperparâmetros, que incluem aspectos como taxa de aprendizado, número de árvores, profundidade máxima e regularização, impactam diretamente no balanceamento entre viés e variância, na mitigação de *overfitting* e na melhora da capacidade preditiva dos modelos. A escolha adequada desses parâmetros não apenas garante maior precisão

preditiva, mas também otimiza o tempo de processamento e a interpretabilidade dos resultados. Na próxima seção, são apresentados os métodos utilizados para ajustar esses hiperparâmetros, bem como os critérios adotados para selecionar as configurações que maximizam a eficácia dos modelos nos diferentes horizontes de previsão. No Apêndice B são apresentados o hiperparâmetros testados em cada modelo.

3.7 Critério de seleção dos melhores hiperparâmetros

Para todos os modelos de aprendizado de máquina, o ajuste dos hiperparâmetros foi realizado com o objetivo de maximizar a área sob a curva ROC (AUC) no conjunto de validação. Inicialmente, foram testadas combinações de hiperparâmetros previamente definidos para cada algoritmo, com base nos trabalhos de Sigrist e Leuenberger (2022) e Jabeur et. al (2021). A análise inicial utilizou o horizonte de 1 período, permitindo identificar os conjuntos de parâmetros mais promissores com base no desempenho do AUC. Essa escolha foi motivada pelo elevado tempo de processamento exigido por alguns modelos em horizontes mais longos.

Posteriormente, os hiperparâmetros foram otimizados para cada horizonte de previsão $j \in \{1, \dots, 12\}$. As Tabelas do Anexo 5.2 apresentam os hiperparâmetros avaliados para cada modelo e os tempos de processamento.

Após a identificação dos melhores hiperparâmetros no conjunto de validação, o modelo com o maior valor de AUC foi selecionado e avaliado no conjunto de testes para todos os horizontes de 1 a 12 trimestres.

A partir dessa configuração, comparações foram realizadas entre o modelo Altman EMS (2005) e Random Forest (melhor modelo), com foco no horizonte de 4 trimestres, utilizando as métricas complementares discutidas na Seção 2.2: (i) o cálculo do spread de crédito; (ii) a medida do custo esperado de classificação incorreta (CECI); e (iii) o impacto no valor da empresa alavancada (VL). Esse procedimento permitiu uma avaliação mais abrangente dos modelos em relação à medida AUC, considerando não apenas sua capacidade discriminatória, mas também suas implicações econômicas e financeiras.

3.8 Considerações sobre viés de sobrevivência nas amostras

A amostra utilizada neste estudo, construída a partir de dados financeiros da plataforma Valor PRO, está sujeita a potenciais vieses de sobrevivência e de seleção, que merecem consideração metodológica. O viés de sobrevivência ocorre porque a base inclui majoritariamente empresas que permaneceram ativas e que, consequentemente, apresentam padrões financeiros menos extremos do que aquelas que desapareceram ou faliram antes do período analisado ou cujos dados não estavam suficientemente disponíveis. Este tipo de viés é amplamente reconhecido na literatura financeira por distorcer inferências estatísticas, levando a uma superestimação da estabilidade e do desempenho das entidades analisadas, conforme demonstrado por Elton, Gruber e Blake (1996) e reforçado em análises históricas de retornos agregados por Jorda et al. (2019).

Adicionalmente, há indícios de um viés de seleção implícito: a entrada de uma empresa na base de dados da Valor PRO, embora não formalmente documentada neste estudo, é possivelmente condicionada a critérios mínimos de relevância econômica, formalização e disponibilidade de informações contábeis. Assim, pequenas empresas, negócios familiares ou aquelas em situação de informalidade tendem a estar sub-representadas ou ausentes, limitando a diversidade de perfis de risco capturada pela amostra.

Esses vieses podem impactar a análise de diversas formas. Primeiramente, podem reduzir a variabilidade observada nos padrões financeiros associados ao colapso empresarial, atenuando a capacidade dos modelos de aprendizado de máquina de identificar sinais precursores de falência em empresas de menor porte ou em trajetórias mais abruptas de deterioração. Em contrapartida, métodos estatísticos tradicionais, como a regressão logística, que se beneficiam de relações médias mais estáveis, podem apresentar desempenho relativamente robusto nesse contexto, como sugerem os resultados encontrados.

De fato, neste estudo, observou-se que a regressão logística apresentou desempenho estatístico muito próximo ao Random Forest, especialmente em horizontes mais curtos, o que pode ser, em parte, atribuído a esse perfil amostral restrito e mais homogêneo. Apesar disso, o Random Forest manteve superioridade em horizontes mais longos, indicando que, mesmo diante de uma base com menor heterogeneidade, sua capacidade de capturar relações complexas e não lineares conferiu vantagens preditivas.

Por fim, é importante ressaltar que tais vieses não comprometem a validade interna das comparações realizadas, uma vez que todos os modelos foram ajustados sobre a mesma base. Contudo, limitam a generalização dos resultados para contextos com maior presença de empresas pequenas, informais ou de setores não representados na base Valor PRO. Essa limitação constitui um ponto relevante para a interpretação dos resultados e para o delineamento de futuras pesquisas, que poderiam buscar bases mais abrangentes ou alternativas metodológicas para mitigar esse viés.

4 RESULTADOS E DISCUSSÃO

Neste capítulo são apresentados os resultados obtidos com os diferentes modelos de classificação de risco de falência. Inicialmente, a precisão preditiva dos modelos foi medida utilizando a métrica AUC. Especificamente, para cada método e horizonte de previsão, foram calculados os valores de AUC no conjunto de testes utilizando o conjunto de hiperparâmetros que maximizaram o AUC no conjunto de validação. Em seguida, o modelo com o melhor resultado em termos de AUC (Random Forest) é comparado com o modelo tradicional de Altman (EMS) nas medidas econômico-financeiras.

Cabe destacar que, ao longo da análise, o Random Forest se consolidou como o modelo de melhor desempenho, apresentando superioridade não apenas nas métricas estatísticas, mas também nas avaliações econômico-financeiras, como o custo esperado de classificação incorreta (CECI), a precificação de crédito e o valor da empresa alavancada. Diante desse desempenho, a comparação central neste estudo foi estabelecida entre o Random Forest, representando as técnicas modernas de aprendizado de máquina, e o modelo clássico de Altman (EMS), amplamente consolidado na literatura sobre previsão de falências.

Embora a Regressão Logística tenha exibido um desempenho competitivo, com AUCs próximas às do Random Forest, especialmente em horizontes mais curtos, por delimitação de escopo, optou-se por não conduzir todas as análises econômicas e financeiras para esse modelo. Tal decisão visou manter o foco na comparação entre o modelo com melhor desempenho global e o modelo mais tradicional, oferecendo assim uma avaliação clara da evolução metodológica e dos ganhos proporcionados pelas técnicas de aprendizado de máquina.

Além disso, ressalta-se que, mesmo com um desempenho ligeiramente inferior ao do Random Forest, a Regressão Logística possui a vantagem de gerar coeficientes interpretáveis, o que facilita a compreensão dos determinantes do risco de falência. Essa característica faz com que, apesar de não ter sido objeto das análises econômicas aprofundadas neste estudo, a Regressão Logística continue sendo uma ferramenta relevante em contextos em que a explicabilidade das decisões é especialmente valorizada.

4.1 Previsão das Probabilidades de Falência de Empresas Individuais

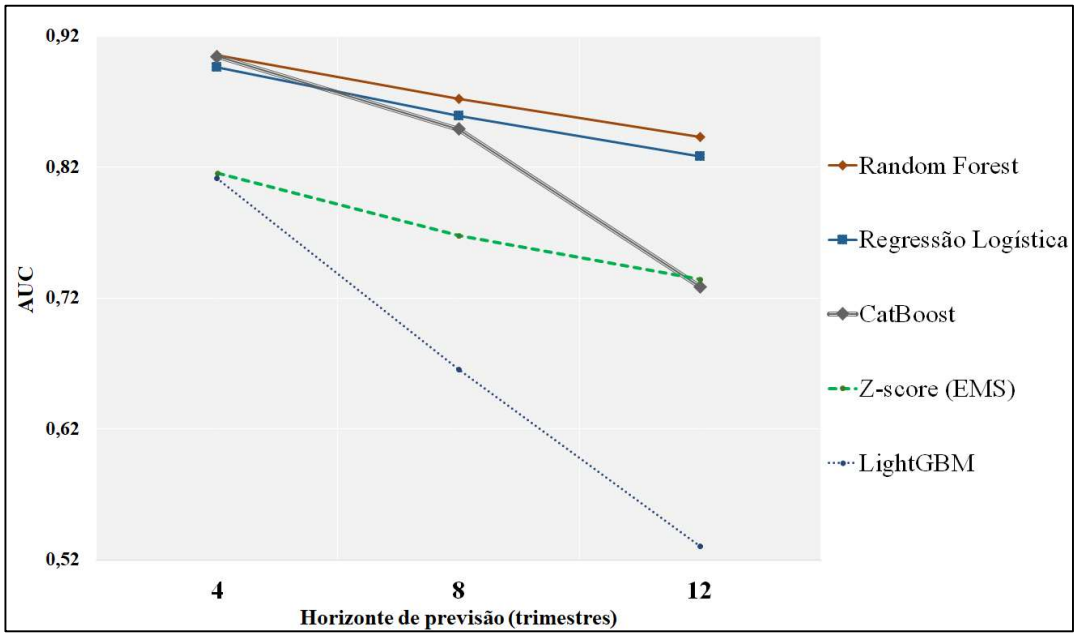
Na Figura 10 observam-se os valores de AUC para três horizontes de previsão (4, 8 e 12 trimestres de antecedência) dos modelos testados. Como esperado, o desempenho dos modelos diminui em termos de AUC à medida que o horizonte analisado se alonga. No Apêndice A, apresentam-se os valores de AUC para os horizontes de 1 a 12 trimestres para cada um dos modelos. Essa tendência de redução do desempenho preditivo pode estar relacionada a diversos fatores. Primeiramente, quanto maior o horizonte de previsão, maior a incerteza associada às variáveis explicativas, o que dificulta a capacidade dos modelos de capturar padrões relevantes e prever eventos de inadimplência. Além disso, alterações na dinâmica econômica e nas condições de mercado ao longo do tempo podem reduzir a relevância das informações históricas para previsões futuras.

Para o horizonte de 4 trimestres, os modelos CatBoost e Random Forest demonstraram desempenhos praticamente idênticos, ambos com os maiores valores de AUC no conjunto de testes. A Regressão Logística apresentou desempenho muito próximo ao desses dois métodos. O modelo de Altman (EMS) obteve um AUC semelhante ao do Light Gradient Boosting Machine (LightGBM), embora ambos tenham se mostrado consideravelmente inferiores aos três primeiros modelos.

No horizonte de 8 trimestres, o Random Forest superou ligeiramente a Regressão Logística, enquanto o CatBoost apresentou um desempenho similar, porém com uma leve redução no AUC em relação aos dois primeiros. O modelo de Altman (EMS) manteve-se inferior aos três métodos, enquanto o LightGBM sofreu uma queda significativa na capacidade discriminatória, apresentando um desempenho muito baixo nesse horizonte.

Para o horizonte mais longo, de 12 trimestres, o Random Forest continuou sendo o modelo com o melhor desempenho, com a capacidade discriminatória da Regressão Logística permanecendo próxima ao do Random Forest. O CatBoost apresentou uma redução maior na AUC, igualando-se ao Altman (EMS). O LightGBM praticamente perdeu sua capacidade discriminatória, com um AUC próximo a 0,5, indicando previsões quase aleatórias nesse horizonte.

Figura 10: AUC dos modelos de previsão de probabilidades de falência.



Observa-se na Figura 10 o AUC vs. horizonte de previsão (em trimestres) para diferentes métodos. 'CatBoost' = gradient boosting com tratamento otimizado para variáveis categóricas, 'RF' = Random Forest, 'LightGBM' = Light Gradient Boosting machine, Regressão Logística, 'Altman EMS' = modelo Z-Score para mercados emergentes de Altman. As linhas representam os valores de AUC nos diferentes horizontes de previsão no conjunto de testes. Fonte: (AUTOR, 2024).

Esses resultados mostram que os métodos de aprendizado de máquina, em especial o Random Forest e o CatBoost, superam o modelo tradicional de Altman (SEM) e o LightGBM em horizontes de previsão mais curtos. Notavelmente, o Random Forest manteve um desempenho superior mesmo em horizontes mais longos, enquanto o LightGBM mostrou-se menos robusto, com desempenho próximo ao aleatório em 12 trimestres de antecedência.

De maneira geral, os resultados mostram que os métodos de aprendizado de máquina superam o modelo Altman (EMS) em termos de AUC, especialmente em horizontes mais curtos e para horizontes mais longos o RF e o Catboost foram mais robustos. Isso demonstra a maior capacidade desses modelos em capturar padrões complexos nos dados financeiros das empresas em horizontes mais longos. Entre os métodos de aprendizado de máquina, o Random Forest e o CatBoost destacaram-se com os melhores desempenhos nos horizontes de 4 e 8 trimestres de antecedência. No entanto, no horizonte de 12 trimestres, o Random Forest manteve sua superioridade, enquanto o CatBoost apresentou uma redução significativa na AUC.

A Regressão Logística, apesar de ser um método estatístico tradicional, mostrou-se competitiva, com desempenho próximo aos melhores modelos de aprendizado de máquina em todos os horizontes analisados. Por outro lado, o LightGBM apresentou desempenho similar ao modelo de Altman nos horizontes mais curtos, mas sofreu uma queda acentuada em horizontes mais longos, evidenciando limitações na sua capacidade preditiva a longo prazo. Esses resultados indicam a importância de considerar a natureza dos dados e o horizonte temporal ao escolher o modelo mais adequado para prever a falência de empresas.

Esses achados ressaltam a importância de considerar o horizonte de previsão ao avaliar modelos de risco de falência. Enquanto alguns modelos complexos de aprendizado de máquina como Catboost oferecem vantagens em horizontes mais curtos, sua eficácia pode diminuir com o tempo.

4.2 Calibração das Probabilidades

A calibração de modelos de previsão de falências avalia o quão bem as probabilidades preditivas refletem as taxas reais de ocorrência dos eventos, Sigrist e Leuenberger (2022), ou seja, dado uma probabilidade prevista p prevista por um modelo de classificação binário, espera-se que o evento real deva ocorrer com essa probabilidade p, Gneiting, Balabdaoui, & Raftery (2007). Métodos de aprendizado de máquina, como random forest e técnicas de boosting, frequentemente apresentam alta precisão discriminatória, mas podem gerar probabilidades preditivas mal calibradas Sigrist e Leuenberger (2022).

Para medir a calibração dos modelos neste estudo, adotou-se o Erro de Calibração Esperado (ECE) Sigrist e Leuenberger (2022) que calcula a diferença média ponderada entre as probabilidades preditivas e as frequências observadas dos eventos, oferecendo uma métrica direta da qualidade das previsões em relação à probabilidade real do evento.

Para calcular o ECE, as probabilidades preditivas foram agrupadas em intervalos e comparado a probabilidade média predita em cada intervalo com a taxa real de falência observada nesse grupo. Matematicamente, o ECE é definido como:

$$ECE = \frac{1}{n}\sum_{i=1}^{B} n_i \, |o_i - \hat{p}_i|$$

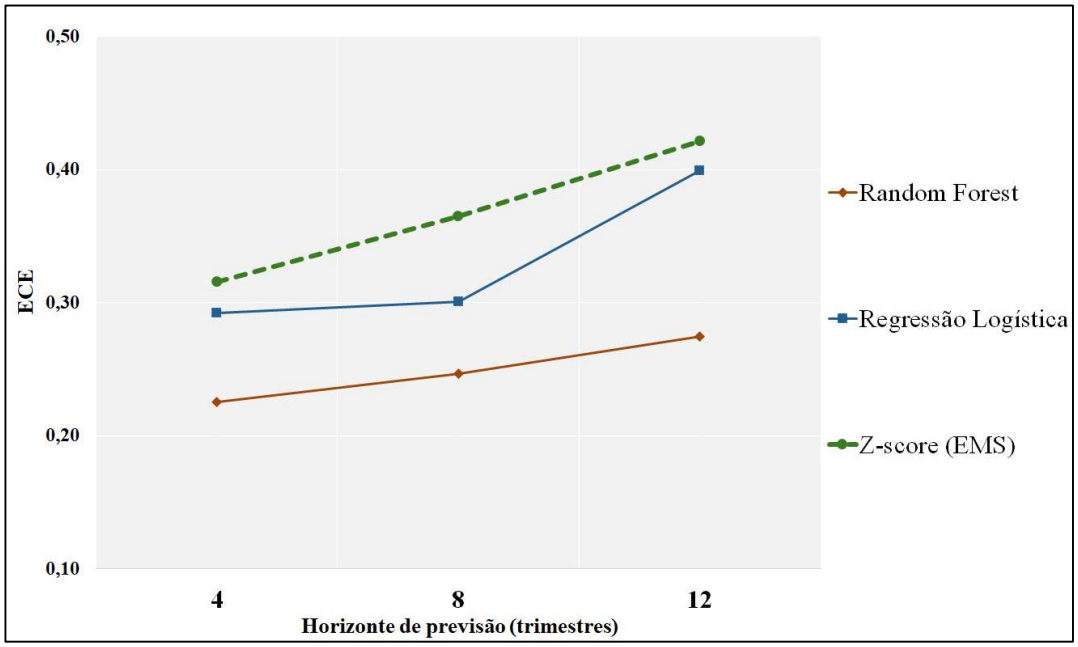
onde:

- n_i é o número total de observações;
- B é o número de intervalos;
- n_i é o número de observações no intervalo i;
- o_i é a taxa observada de falência no intervalo i;
- $\hat{p}_i$ é a probabilidade preditiva média no intervalo i.

Essa metodologia foi aplicada aos modelos que apresentaram maior AUC no conjunto de testes: Random Forest. Regressão Logística e Altman (EMS), nos horizontes de previsão de 4, 8 e 12 trimestres criando 10 intervalos entre 0 e 1 espaçados em 0,1.

Os resultados indicaram que o Random Forest apresentou o menor ECE em todos os horizontes de previsão, demonstrando uma calibração superior em relação aos outros modelos, como pode ser observado na Figura 11. A Regressão Logística exibiu um ECE menor do que o modelo de Altman (EMS) nos horizontes de 4 e 8 trimestres, mas no horizonte de 12 trimestres o ECE aumentou a ponto de aproximando-se do ECE do modelo de Altman (EMS). Isso sugere que a capacidade da Regressão Logística em produzir probabilidades bem calibradas diminui em previsões de longo prazo, enquanto o Random Forest mantém sua robustez mesmo em horizontes mais longos.

Figura 11: Erro de calibração esperada.



ECE nos horizontes de 4, 8 e 12 trimestres de antecedência, para os modelos Random Forest, Regressão Logísticas e Altman (EMS). Fonte: (AUTOR, 2024).

Uma abordagem para melhorar a calibração das probabilidades preditivas é o uso de funções monotônicas Sigrist e Leuenberger (2022), como o Platt Scaling, Platt et al., (1999) ou a regressão isotônica Zadrozny & Elkan, (2001). Essas técnicas ajustam as probabilidades preditivas originais através de transformações que preservam a ordem dos riscos (monotonicidade), mas alteram a escala para melhor alinhar as previsões com as frequências observadas dos eventos.

Outra abordagem para lidar com inconsistências nas probabilidades preditivas foi apresentada por Burgt (2007), que propôs um método de suavização para assegurar a monotonicidade das probabilidades em categorias de risco. Esse método é útil em situações em que a segmentação em grupos esparsos pode levar a probabilidades incoerentes, como uma probabilidade de falência mais baixa atribuída a uma categoria de risco mais alta. A técnica de Burgt (2007) ajusta as probabilidades observadas de modo que estas aumentem ou diminuam de forma contínua e consistente, garantindo que os modelos reflitam adequadamente a relação entre risco e categoria.

Para realizar a calibração das probabilidades preditivas dos modelos de previsão de falência nesse estudo, foi utilizado o método de regressão isotônica implementado com o pacote *isotone* no software R, Mair et al., 2023. Esse procedimento foi aplicado aos modelos Random Forest (modelo com maior AUC) e Altman EMS (modelo mais tradicional) para o horizonte de 4 trimestres, a fim de possibilitar comparações das métricas CECI, precificação de crédito e impacto sobre o valor da empresa alavancada V^L.

A calibração isotônica foi realizada agrupando as probabilidades preditivas em intervalos de 0,001, calculando a média observada de falência em cada intervalo. Em seguida, ajustou-se as probabilidades médias por intervalo para garantir que fossem monotônicas crescentes, de modo a preservar a relação esperada entre probabilidades maiores e maior risco de falência. Esse ajuste não paramétrico é útil para suavizar inconsistências nos dados, especialmente em amostras desbalanceadas ou em situações em que os modelos iniciais apresentem problemas de calibração Sigrist e Leuenberger (2022).

No modelo Random Forest e no Altman (SEM), a calibração resultou em menores valores de ECE, conforme mostrado na Tabela 8.

Tabela 8: Erro de calibração esperado antes de depois da calibração das probabilidades.

	ECE	
Modelo	Antes calibração	Após calibração isotônica
Altman (EMS)	3,6E-01	3,0E-18
Random Forest	2,5E-01	1,6E-18

ECE para os modelos Altman (EMS) e Random Forest para o horizonte de 4 trimestres de antecedência.
Fonte: (AUTOR, 2024),

Esse ajuste é importante em aplicações como a análise de risco de crédito, precificação de empréstimo e gestão de portifólios de dívidas, pois garante que as probabilidades preditivas e as distribuições de perdas esperadas reflitam adequadamente o risco real Sigrist e Leuenberger (2022).

Usando as probabilidades calibradas, nos próximos tópicos são discutidos os desempenhos dos modelos em termos de CECI, precificação de crédito e Valor da Empresa Alavancada (VL), conectando as previsões probabilísticas às suas implicações econômicas e financeiras.

4.3 Custo esperado de classificação incorreta (CECI)

O procedimento adotado para comparar os custos esperados de classificação incorreta (CECI), dos modelos Random Forest (com melhor AUC) e Altman (SEM) com 4 trimestres de antecedência seguiu o framework proposto por Jardin (2016). Esse trabalho enfatiza a importância de ajustar os valores de corte dos modelos de classificação levando em consideração os custos assimétricos associados aos erros de classificação tipo I e tipo II.

A abordagem descrita por Jardin permite adaptar os modelos ao contexto de aversão ao risco dos tomadores de decisão, em um cenário de alta ou baixa aversão ao risco.

A fórmula utilizada para o cálculo é:

$$Custo\ esperado\ de\ classificação\ incorreta\ (CECI) = \frac{c_1 . p_1 . e_1}{n_1} + \frac{c_2 . p_2 . e_2}{n_2}$$

Onde:

- c_1 e c_2: custos dos erros Tipo I e Tipo II, respectivamente;

- e_1 e e_2: taxas de erros Tipo I e Tipo II, respectivamente;
- n_1 e n_2: número de empresas solventes e falidas na amostra, respectivamente;
- p_1 e p_2: probabilidades a priori de não falência e de falência, respectivamente.

No caso de bancos que priorizam a minimização de perdas financeiras significativas decorrentes de inadimplências (erro tipo I), pode-se adotar uma relação c_1/c_2 maior, como 100/1, atribuindo um peso maior aos custos de classificar erroneamente uma empresa falida como solvente. Por outro lado, em um cenário voltado à maximização de oportunidades de crédito, como em empresas que desejam expandir suas carteiras mesmo assumindo maior risco de inadimplência, uma relação como 10/1, refletindo uma menor penalização para os erros do tipo I.

Foram comparados os CECI dos dois métodos, Random Forest e Altman EMS, utilizando seis cenários distintos para a relação c_1/c_2, conforme Jardin (2016); cenário 1, $c_1 = c_2$ onde custo de classificar uma empresa com dificuldades financeiras como solvente é igual ao custo de classificar uma empresa solvente como falida. Cenário 2: $c_1 = 20c_2$ no qual o custo de classificar uma empresa com dificuldades financeiras como solvente é vinte vezes mais alto que custo de classificar uma empresa solvente como falida. Cenário 3: $c_1 = 40c_2$, onde o custo de classificar uma empresa com dificuldades financeiras como solvente é quarenta vezes mais alto que custo de classificar uma empresa solvente como falida, cenário 4: $c_1 = 60c_2$, cenário 5: $c_1 = 80c_2$, e cenário 6: $c_1 = 100c_2$

A análise dos resultados apresentados na Tabela 9 evidencia que o Random Forest foi mais eficiente que o modelo Altman (EMS) em minimizar os custos esperados de classificação incorreta (CECI) na maioria dos cenários analisados. Para a relação $c_1 = c_2$, o modelo Random Forest apresentou um CECI de 24,40, ligeiramente inferior ao valor de 24,98 obtido pelo modelo EA.

À medida que a relação c_1/c_2 aumenta, refletindo uma maior penalização para os erros tipo I (classificar empresas falidas como solventes), o Random Forest continua apresentando menores CECI's em quase todos os cenários. Para $c_1 = 40c_2$, o Random Forest registra um CECI de 348,96, enquanto o EA apresenta um valor de 511,50, que demonstra a superioridade do Random Forest em situações em que os custos de erros tipo I são elevados. Além disso, para $c_1 = 80c_2$, o Random Forest obtém um CECI de 47,42, comparado aos 140,36 do modelo EA, reforçando o desempenho mais consistente do Random Forest em cenários mais extremos

Tabela 9: Custo esperado de classificação incorreta para EMS e Random Forest para 4 trimestres.

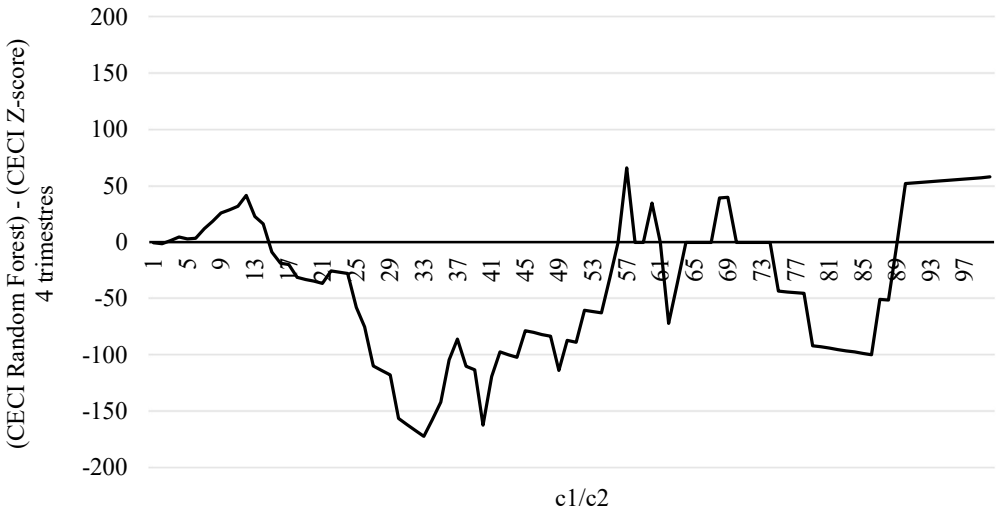
Modelo	c1/c2	CECI	Limiar	Acurácia	Erro Tipo I	Erro Tipo II	Sensibilidade	Precisão	Especificidade	p-valor (Acurácia)
Altman EMS	1	24,98	0,5000	0,9989	0,9773	0,0000	0,0227	0,5000	1,0000	-
	20	348,66	0,0476	0,9962	0,6818	0,0031	0,3182	0,1045	0,9969	-
	40	511,50	0,0244	0,9895	0,5000	0,0099	0,5000	0,0538	0,9901	-
	60	244,76	0,0164	0,9704	0,1591	0,0294	0,8409	0,0312	0,9706	-
	80	140,36	0,0123	0,9635	0,0682	0,0364	0,9318	0,0280	0,9636	-
	100	1,11	0,0099	0,9568	0,0000	0,0432	1,0000	0,0254	0,9568	-
Random Forest	1	24,40	0,5000	0,9989	0,9545	0,0001	0,0455	0,5000	0,9999	1,0000
	20	313,85	0,0476	0,9941	0,6136	0,0052	0,3864	0,0769	0,9948	0,0000
	40	348,96	0,0244	0,9848	0,3409	0,0148	0,6591	0,0476	0,9852	0,0000
	60	279,49	0,0164	0,9755	0,1818	0,0244	0,8182	0,0364	0,9756	0,0000
	80	47,42	0,0123	0,9632	0,0227	0,0369	0,9773	0,0290	0,9631	0,7789
	100	59,04	0,0099	0,9632	0,0227	0,0369	0,9773	0,0290	0,9631	0,0000

Os valores do CECI foram multiplicados por 1 milhão para facilitar a interpretação e comparar os custos esperados em uma escala mais prática e significativa no contexto econômico da análise. Foi aplicado o teste de McNemar, detalhado no Anexo A, para avaliar as diferenças de acurácia entre os dois modelos. O teste indicou significância estatística ($p < 0{,}001$) para a maioria dos cenários, exceto para $\frac{c_1}{c_2} = 1, e \frac{c_1}{c_2} = 80$. Isso confirma que o Random Forest superou o modelo EA na maioria dos cenários em termos de minimização do CECI e desempenho geral de classificação.

No entanto, é importante observar que, para $c_1 = 60c_2$ e $c_1 = 100c_2$, o modelo EA apresenta um CECI menor do que o Random Forest. Esse resultado sugere que, em outros cenários de custos maiores associados a erros tipo I, o modelo Altman (EMS) pode ser mais eficiente. Para conclusões mais robustas, seria recomendável avaliar ambos os modelos em diferentes horizontes de previsão, considerando prazos superiores a quatro trimestres, que foi o limite do escopo desta análise. Essa extensão permitiria compreender melhor o comportamento dos modelos em contextos de diferentes prazos, onde as características dos dados e as relações entre variáveis podem mudar.

Para ilustrar essa limitação, observa-se na Figura 12 um gráfico da diferença entre o CECI do Random Forest e o CECI do modelo Z-score, com o eixo y representando a diferença (CECI Random Forest - CECI Z-score) e o eixo x representando os valores de c_1/c_2, variando de 1 a 100. Nesse gráfico, valores abaixo de zero indicam que o CECI do Random Forest foi menor do que o CECI do Z-score, enquanto valores acima de zero indicam o oposto. Observa-se que o Random Forest superou o Z-score em 73% dos cenários analisados.

Figura 12: CECI para 100 cenários da relação do custo do erro tipo I e tipo II



Fonte: AUTOR (2024)

Embora o gráfico indique uma vantagem geral do Random Forest, a magnitude das diferenças não apresenta um padrão claro ou limites mínimos e máximos consistentes. Isso sugere que conclusões mais robustas poderiam ser obtidas ao ampliar a análise para incluir horizontes de previsão mais longos, permitindo explorar como a diferença no desempenho dos modelos evolui em diferentes contextos temporais e condições de custos.

Considerando o escopo dos 6 cenários conforme o framework de Jardin (2016), além do CECI, métricas baseadas na matriz de confusão ajudam a entender como os modelos se comportam em diferentes cenários. Para valores de c_1/c_2 mais baixos, o Random Forest mantém uma sensibilidade elevada (chegando a 97,73% para $c_1/c_2 = 80$ e $c_1/c_2 = 100$, eliminando quase completamente os erros tipo II, mas sacrificando a especificidade (96,31%). Esse *trade-off* é esperado, dado o aumento do peso atribuído às falências na minimização do CECI. Em contraste, o modelo Altman (EMS) apresenta uma sensibilidade mais baixa em cenários com penalizações mais altas para erros tipo I, como $c_1/c_2 = 60$ e $c_1/c_2 = 80$ sugerindo uma maior dificuldade em ajustar-se a esses custos.

Os resultados do CECI e das métricas de desempenho baseadas na matriz de confusão demonstram que o Random Forest oferece uma maior capacidade de adaptação a cenários variados, permitindo decisões de crédito mais alinhadas ao risco real dos tomadores. Ao minimizar os erros tipo II (classificar solventes como falidas), especialmente em cenários com custos elevados associados à classificação de empresas insolventes como solventes, o Random Forest permite uma precificação mais precisa e contribui para uma alocação

mais eficiente de recursos financeiros. Em contraste, o modelo EA, embora ainda robusto para o CECI em alguns cenários, é mais sensível aos custos atribuídos aos erros, resultando em um desempenho menos consistente em termos de sensibilidade.

4.4 Rentabilidade de modelos de previsão de falências

Para avaliar a rentabilidade dos modelos de previsão de falências, foi realizada uma comparação entre o Random Forest e o modelo Altman (EMS), utilizando probabilidades calibradas para previsões com 4 trimestres de antecedência. A metodologia seguiu a lógica de competição pareada desenvolvida por Blöchlinger e Leippold (2006), que simula um mercado de crédito competitivo em que tomadores escolhem o modelo com menor spread de crédito. Essa abordagem foi também utilizada em estudos posteriores, como Agarwal e Taffler (2008), Agarwal e Bauer (2014) e Lohmann e Möllenhoff1 (2022), sendo relevante para conectar a capacidade preditiva dos modelos a métricas econômicas, como lucro, market share e retorno sobre ativos. Cada empresa na amostra foi avaliada em termos de precificação de crédito utilizando a fórmula de spread, definida por:

$$R\ (t) = \frac{P(Y = 1|S = t)}{P(Y = 0|S = t)}.LGD + k$$

onde k foi fixado como o retorno mínimo de 0,3%, e o limite máximo para o spread foi estabelecido em 10%. Foram considerados três cenários de perdas em caso de inadimplência (LGD): 15%, 45% e 75%.

Os modelos competem simulando um mercado de crédito com $1.000,0 unidades monetárias disponíveis por trimestre. A empresa que apresentasse o menor spread era atribuída ao modelo correspondente. Em casos de empate no spread, os clientes eram divididos igualmente entre os dois modelos, com cada um cedendo 50% do crédito. Para empréstimos que não foram cedidos por um modelo, mas que foram concedidos pelo outro a empresas que não faliram, foi contabilizado um custo de oportunidade no cálculo do lucro, utilizando a taxa mínima.

Os resultados foram avaliados com base em quatro medidas de desempenho: (i) participação de mercado (market share), (ii) margem de lucro (lucro/receita), (iii) retorno sobre ativos (ROA: lucro/ativos emprestados) e (iv) retorno sobre ativos ponderados pelo risco (RWA: lucro/ativos ponderados pelo risco). O cálculo do ativo ponderado pelo risco

seguiu as diretrizes do framework Basiléia III (European Union, 2013, Anexo B), como descrito por Lohmann e Möllenhoff1 (2022), assumindo uma maturidade de crédito de 2,5 anos. A análise considerou os dados reais de falência ou não falência para calcular os resultados efetivos de cada medida e identificar a eficiência econômica dos modelos. Os empréstimos não cedidos para um modelo, mas cedido pelo outro para empresas que não faliram foram consideradas como custo de oportunidade no cálculo do lucro à taxa mínima k. Essa abordagem permitiu avaliar o desempenho dos modelos em um ambiente competitivo de crédito, conectando previsões probabilísticas às suas implicações econômicas reais. Os resultados podem ser observados na Tabela 10 e evidenciam a superioridade do modelo Random Forest em relação ao Altman (EMS) em todos os cenários analisados de LGD.

Tabela 10: Análise comparativa de modelos considerando a política de concessão de crédito e a rentabilidade (conjunto de treinamento).

Resultados da precificação de crédito	LGD = 15%		LGD = 45%		LGD = 75%	
	Altman (EMS)	Random Forest	Altman (EMS)	Random Forest	Altman (EMS)	Random Forest
Market Share (%)	49,94	50,06	49,94	50,06	49,94	50,06
Lucro ($)	18,10	18,51	15,71	16,69	13,33	14,87
Margem de lucro (%)	85,72	87,85	73,23	78,75	61,12	69,75
Retorno sobre ativos (ROA-%)	0,2589	0,2641	0,2248	0,2382	0,1906	0,2122
Retorno ativos ponderados (RWA-%)	0,145	0,147	0,042	0,044	0,021	0,024

Em termos de market share, ambos os modelos obtiveram uma divisão quase equitativa do mercado (49,94% para o Altman (EMS) e 50,06% para o Random Forest), sugerindo que, apesar das diferenças nas probabilidades preditivas e spreads calculados, os modelos competem de maneira equilibrada em atrair tomadores de crédito.

Entretanto, as métricas financeiras destacam uma vantagem econômica do Random Forest. Em todos os cenários de LGD, o Random Forest gerou maior margem de lucro, com a diferença aumentando conforme os custos de inadimplência se elevam. Por exemplo, no cenário de LGD = 75%, o lucro do Random Forest foi 8,63 pontos percentuais maior que o do Altman (EMS).

Além disso, nos indicadores de retorno sobre ativos (ROA) e retorno sobre ativos ponderados pelo risco (RWA), o Random Forest apresentou melhor desempenho em todos os cenários. Para o LGD mais elevado (75%), o ROA do Random Forest foi 0,002 pontos percentuais superior ao do Altman (EMS), enquanto o RWA foi 0,002 p.p. maior no mesmo cenário. Esses resultados indicam que o Random Forest oferece não apenas maior lucratividade, mas também maior resiliência em cenários de risco elevado.

Esses resultados corroboram com os achados de Agarwal e Taffler (2014), no qual enfatizaram que diferenças mesmo pequenas na acurácia preditiva podem gerar impactos significativos nas métricas econômicas, especialmente em cenários com custos assimétricos de classificação. Esse ponto é consistente com os resultados observados, nos quais o Random Forest mostrou-se mais robusto em termos de AUC lucro e margem, mesmo em cenários de alto LGD.

4.5 Impacto dos modelos de previsão no valor da empresa alavancada

Para avaliar o impacto dos modelos de previsão de falências no valor da empresa alavancada (V^L), utilizou-se a estrutura conceitual baseada na Teoria do *Trade-off*. Essa teoria diz que a empresa equilibra os benefícios fiscais do endividamento com os custos esperados de falência DeMarzo (2010). Foram comparados os modelos Random Forest e Altman (EMS) no horizonte de 4 trimestres em usando suas previsões calibradas. Essa avaliação considerou uma observação de uma empresa considerada mediana pelo valor do seu ativo em relação à amostra desse estudo. A empresa foi denominada como "X". O cálculo seguiu as fórmulas a seguir:

4.5.1 Cálculo do valor da empresa alavancada

$$V^L_{(x,j,w)} = V^U_{(x,j)} + VP_{(x,j,w)}(BF) - VP_{(x,j,w)}(DF)$$

Onde:

- $V^L_{(x\ j,w)}$: valor da empresa alavancada x, no trimestre j, calculada pelo modelo w;
- $V^U_{(x,j)}$: valor da empresa desalavancada x, no trimestre j.
- $VP_{(x,j,w)}(BF)$: valor presente dos benefícios fiscais das despesas com juros, dado o modelo w para a empresa x;
- $VP_{(x,j,w)}(DF)$: valor presente dos custos de dificuldades financeiras, dado o modelo w para a empresa x.

4.5.2 Cálculo do valor da empresa desalavancada

O valor da empresa desalavancada $V^U_{(x,j)}$ foi estimado com base no fluxo de caixa operacional de uma observação, sem projeções futuras seguindo o objetivo de uma abordagem simplificada e comparativa:

$$V^U_{(x,j)} = \frac{FC_{(x,j)}}{1 + k_e}$$

Onde:

- $FC_{(i,j)}$:fluxo de caixa da empresa x no trimestre j;
- k_e: custo de capital próprio, definido hipoteticamente como 20% para os dois modelos.

4.5.3 Valor Presente dos benefícios fiscais da dívida

Os benefícios fiscais das despesas com juros foram calculados considerando que os juros pagos sobre a dívida são dedutíveis para fins fiscais, proporcionando uma economia tributária. O valor presente desses benefícios é dado por:

$$VP_{i,j,w}(BF) = \frac{\tau \cdot k_{d_{(x,j,w)}} \cdot D_{(x,j)}}{1 + k_e}$$

Onde:

- τ: alíquota de imposto de renda, fixada em 34%;
- $k_{d_{(x,j,w)}}$: custo da dívida da empresa x no trimestre j, com base no modelo w;
- $D_{(i,j)}$: saldo de dívida no balanço da empresa x no trimestre j;
- k_e: custo de capital próprio, definido hipoteticamente como 20% para os dois modelos.

4.5.4 Cálculo do Custo da Dívida

O custo da dívida é calculado incorporando o risco de crédito da empresa, influenciado pela probabilidade de default estimada pelos modelos, conforme metodologia descrida na seção 4.4 e dado por:

$$k_{d_{(x,j,w)}} = \frac{P_{(i,j,w)}(Y = 1|S = t)}{P_{(i,j,w)}(Y = 0|S = t)}.LGD + k$$

- k: spread mínimo exigido pelo credor, fixado em 0,3% conforme metodologia explicada na Seção 4.4 sobre precificação de créditos;
- $P_{(x,j,w)}(Y = 1|S = t)$ é a probabilidade de default da empresa x dado o score t do modelo w;
- $P_{i,j,w}(Y = 0|S = t)$ é a probabilidade da empresa x permanecer solvente dado o score t do modelo w;
- LGD: perda dada a inadimplência, simulada para os valores de 15%, 45% e 75%.

4.5.5 Valor presente dos custos de dificuldades financeiras

O valor presente dos custos de dificuldades financeiras foi calculado considerando a existência dos custos diretos e indiretos de falência DeMarzo (2010). Os custos diretos incluem despesas administrativas relacionadas ao processo falimentar, como honorários advocatícios, taxas judiciais e pagamentos a auditores e peritos. Os custos indiretos estão associados a oportunidades perdidas devido à falência, como dificuldade de acesso a financiamento, queda nas receitas, desvalorização da empresa e demissões. Esses últimos, além de serem mais amplos, apresentam maior complexidade para identificação e mensuração Custódio et.al (2023).

Para a avaliação empírica utilizou-se os resultados de Jupetipe et. al (2017), esse estudo avaliou 131 casos de falência e recuperação judicial de empresas brasileiras, no qual observaram que os custos diretos representavam em média 32,5% do ativo inicial da empresa falida e custos indiretos de 47%, considerando tanto os processos de falência quanto os de recuperação judicial. Assumiu-se essa premissa de custos a todas as empresas da amostra, independentemente de suas características individuais.

O valor presente $VP_{(x,j,w)}(DF)$ dos custos de dificuldades financeiras foi calculado como:

$$VP_{(x,j,w)}(DF) = \frac{P_{(x,j,w)}(Y = 1|S = t) \cdot (CD_{\%AT} + CI_{\%AT}) \cdot A_{(x,j-1)}}{1 + k_e}$$

Onde,

- $P_{(x,j,w)}(Y = 1|S = t)$ é a probabilidade de default da empresa x dado o score t do modelo w.

- $CD_{\%AT}$: é o custo direto de falência como percentual do ativo da empresa;
- $CI_{\%AT}$: é o custo indireto de falência como percentual do ativo da empresa;
- $A_{(i,j)}$ ativo total médio da empresa x no trimestre $j-1$.

4.5.6 Comparação do valor da empresa alavancada para diferentes modelos

As diferenças observadas nos resultados do Valor da empresa alavancada podem ser observadas na Tabela 11 e decorrem fundamentalmente das diferentes probabilidades de default (PD) atribuídas pelos modelos de previsão (EMS e RF) e das implicações que essas probabilidades exercem sobre o custo da dívida e, consequentemente, sobre o valor alavancado da empresa.

Variações nas probabilidades de default alteram a percepção de risco por parte dos credores, refletindo-se no custo da dívida e, por extensão, na magnitude dos benefícios fiscais e dos custos de dificuldades financeiras.

Tabela 11: Valor da empresa alavancada dado pelos modelos Random Forest e Altman (EMS)

Variável Fixa	Valor observado	Variável calculada por Modelo	Valor calculado EMS	Valor calculado RF
$FC_{(x,j)}$	\$20.434	$P_{(x,j,w)}(Y = 1\|S = t)$	1,10%	1,51%
τ	34%	$k_{d_{(x,j,w)}}$	1,411%	1,833%
k_e	20%	$V^U_{(x,j)}$	\$17.029	\$17.029
k	0,30%	$VP_{(i,j,w)}(BF)$	\$240	\$312
$CD_{\%AT}$	33%	$VP_{(x,j,w)}(DF)$	\$6.491	\$8.917
$CI_{\%AT}$	47%	$V^L_{(x,j,w)}$	\$10.777	\$8.423
$A_{(x,j)}$	\$891.606			
$D_{(x,j)}$	\$59.997			

Valores apresentados em R$ mil. Fonte: (AUTOR, 2024).

Quando um modelo atribui uma probabilidade de default mais elevada, o risco de insolvência percebido pelo mercado de crédito é maior, o que se traduz em um acréscimo no custo da dívida k_d. O aumento do k_d implica em duas dinâmicas simultâneas. Por um lado, o maior custo da dívida, ao elevar a base de cálculo do benefício fiscal, tende a ampliar o valor presente dos benefícios tributários oriundos da dedutibilidade dos juros. Por outro lado, essa mesma elevação na percepção de risco amplia a probabilidade e a intensidade dos custos de dificuldades financeiras.

Segundo a teoria do *trade-off*, a relação entre esses efeitos não é linear. Embora um custo de dívida mais alto possa elevar os benefícios fiscais, o ganho resultante pode ser mais do que compensado pelo crescimento dos custos esperados de dificuldades financeiras, notadamente quando o risco de default atinge níveis mais expressivos Demarzo (2010). Ou seja, a elevação na PD estimada por um dado modelo pode aumentar o valor presente dos benefícios fiscais a um custo ainda maior, na forma de custos esperados de falência maiores. Como resultado, o valor da empresa alavancada $V^L_{(x,j,w)}$ pode diminuir, na medida em que a deterioração da qualidade do crédito supera os ganhos tributários adicionais.

Observa-se na Figura 13 a dinâmica simultânea da teoria do *trade-off* da estrutura de capital. Nessa figura são apresentados dois gráficos com duas curvas cada um para a empresa "X". O gráfico superior mostra o $\boldsymbol{V^L_{(x,j,w)}}$ para diferentes cenários de endividamento considerando o cálculo, enquanto o gráfico da parte inferior mostra o custo

médio ponderado de capital para a mesma empresa “X” em vários níveis de endividamento relativo dado pelo $\frac{D}{D+PL}$, onde D é o saldo de dívida e PL o patrimônio líquido da empresa.

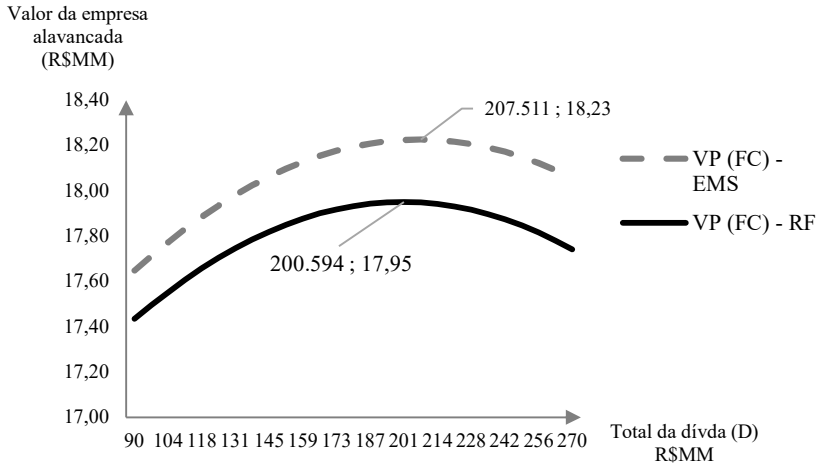
A análise comparativa entre diferentes modelos de previsão - no caso, um modelo que atribui probabilidades de default mais conservadoras (ou maiores), nesse caso o modelo RF e outro que as atribui probabilidades menores (nesse exemplo o modelo Altman EMS) - permite avaliar a sensibilidade do valor da empresa às estimativas de risco. Um modelo mais "pessimista" (RF) em termos de PD tende a precificar o endividamento em um patamar mais elevado, implicando em maior spread exigido pelo credor e, assim, em um custo de capital de terceiros superior. Ao mesmo tempo, um modelo mais "otimista" (com PD menor, nesse exemplo o EMS) poderia subestimar o risco, levando a uma redução no custo da dívida e, portanto, a benefícios fiscais mais modestos, mas também a custos esperados de dificuldades financeiras relativamente menores. Nesse sentido, o valor alavancado da empresa refletirá a qualidade informacional do modelo de previsão adotado: um modelo com estimativas mais acuradas do risco de default tende a fornecer uma visão mais equilibrada sobre o *trade-off* entre benefícios fiscais e custos de falência, contribuindo para uma estimação mais confiável do valor da firma.

Essas considerações assumem mais importância em contextos em que a seleção da estrutura de capital depende da percepção de risco. Empresas que buscam maximizar seu valor podem estar tentadas a aumentar seu grau de endividamento se os modelos de previsão sugerirem riscos mais baixos do que realmente são. Entretanto, a médio e longo prazo, na medida em que os custos de falência não previstos adequadamente pelo modelo passam a ser mais transparentes para os credores e o mercado as novas informações podem comprometer o valor real da empresa. Por outro lado, modelos que superestimam a probabilidade de default podem desencorajar níveis adequados de endividamento, levando a uma subutilização de benefícios fiscais e possivelmente reduzindo o valor da firma.

Assim, a análise econômica das diferenças entre os resultados gerados por distintos modelos de previsão de default destaca a relevância da acurácia na estimativa do risco para a determinação do valor alavancado da empresa. Modelos que capturam melhor a relação entre probabilidade de default e custos de dificuldades financeiras podem fornecer orientações mais consistentes para a definição da estrutura de capital, enquanto modelos

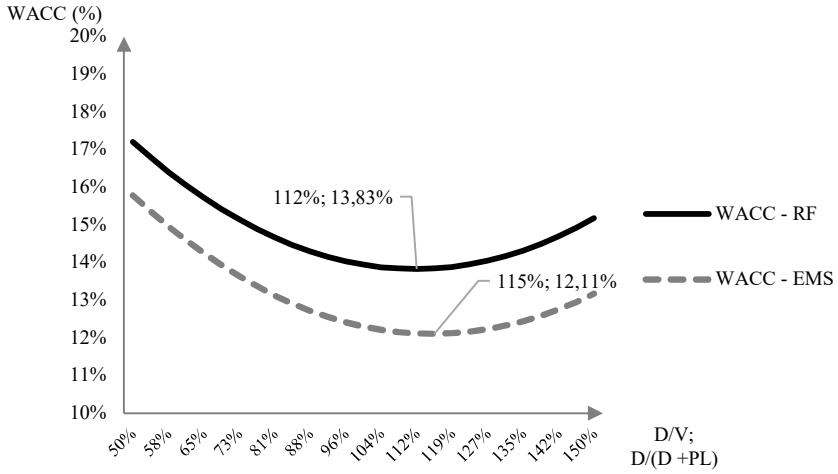
menos precisos podem conduzir a conclusões distorcidas sobre o nível ótimo de endividamento.

Figura 13: Valor alavancado da empresa para dois modelos de previsão de falência



Fonte: (AUTOR, 2024). Para construção das curvas a fórmula dos juros da dívida foi ajustado a fim de refletir a relação não linear entre o custo da dívida o nível de alavancagem e ilustrar os efeitos da teoria do trade-off. WACC – RF= é o custo ponderado de capital considerando as probabilidades atribuídas pelo modelo Random Forest e WACC – EMS pelo modelo Altman (EMS)

Figura 14: Custo ponderado de capital para dois modelos de previsão de falência



Fonte: (AUTOR, 2024). Para construção das curvas a fórmula dos juros da dívida foi ajustado a fim de refletir a relação não linear entre o custo da dívida o nível de alavancagem e ilustrar os efeitos da teoria do trade-off. WACC – RF= é o custo ponderado de capital considerando as probabilidades atribuídas pelo modelo Random Forest e WACC – EMS pelo modelo Altman (EMS).

5 CONSIDERAÇÕES FINAIS

Este estudo conduziu uma análise do desempenho dos modelos de aprendizagem de máquina na previsão de falências corporativas no Brasil, destacando sua superioridade em relação a métodos tradicionais, como o Z-Score de Altman (EMS), especialmente em horizontes temporais mais longos e em diversas dimensões econômicas. Os resultados destacaram a superioridade dos modelos baseados em aprendizado de máquina, particularmente o Random Forest, na precisão preditiva (AUC). O Random Forest demonstrou-se o modelo mais robusto em horizontes mais longos, evidenciando sua capacidade de capturar relações complexas nos dados financeiros, mesmo em cenários de de maior incerteza e variabilidade.

A principal contribuição deste trabalho é demonstrar como avanços em aprendizado de máquina podem transformar decisões estratégicas em contextos financeiros. Os resultados revelaram que o Random Forest não apenas minimiza os custos esperados de classificação incorreta (CECI), mas também ajusta spreads de crédito de maneira mais alinhada aos perfis de risco das empresas. Ademais, a precisão preditiva mostrou-se crucial para a definição da estrutura ótima de capital, impactando diretamente a tomada de decisões de empresas e credores. Esses achados reforçam a relevância de adotar técnicas modernas de aprendizado de máquina em mercados emergentes como o Brasil, mesmo diante da escassez de dados financeiros em comparação com os volumes amplos utilizados em estudos de mercados desenvolvidos, como os Estados Unidos. Comparando com o trabalho de Sigrist e Leuenberger (2022), que analisou 2.824.412 observações de empresa-mês em um conjunto de 20.235 empresas distintas, esse estudo, embora limitado a 268.685 observações ao longo dos períodos analisados, oferece contribuições relevantes ao explorar o contexto brasileiro, caracterizado por dados mais escassos. Essa diferença ressalta a importância de desenvolver metodologias adaptadas a mercados emergentes, onde as restrições informacionais são mais recorrentes. Essa abordagem permite mitigar riscos financeiros de forma adaptada à realidade local e contribuir para a maximização do valor corporativo em cenários de maior restrição informacional.

Adicionalmente, este trabalho avança ao integrar dimensões econômicas às avaliações estatísticas, ampliando o entendimento sobre a interação entre probabilidade de default, custos de dificuldades financeiras e benefícios fiscais.

No entanto, há limitações que devem ser consideradas, como sugestões para pesquisas futuras como a ausência de variáveis macroeconômicas e de análises mais detalhadas para os demais horizontes em termo econômicos e financeiro. A inclusão de elementos como taxas de juros, câmbio e condições de mercado pode aprimorar a precisão e a relevância dos modelos no contexto brasileiro. Essas iniciativas poderiam oferecer uma compreensão ainda mais abrangente sobre o papel dos modelos de previsão de falência na gestão de riscos financeiros e na formulação de estratégias robustas para mercados emergentes.

Adicionalmente, é importante reconhecer que a composição da base de dados utilizada neste estudo, centrada em empresas com dados contábeis disponíveis e contínuos, pode refletir um viés de sobrevivência. Esse viés, embora não comprometa a comparação entre modelos, pode atenuar diferenças entre abordagens mais simples e mais complexas, ao reduzir a presença de padrões extremos de deterioração financeira. Portanto, sua consideração é fundamental na extrapolação dos resultados e na construção de modelos ainda mais robustos.

Destaca-se também que a amostra analisada é predominantemente composta por empresas de médio e grande porte, com baixa representatividade das micro e pequenas empresas — justamente o segmento que, conforme evidenciado pelos dados mais recentes da Serasa Experian, apresentou o maior crescimento proporcional nos pedidos de recuperação judicial no Brasil. Essa composição, embora assegure a qualidade e consistência das informações utilizadas, constitui uma limitação metodológica importante, restringindo a generalização dos achados para o universo das pequenas e médias empresas. Assim, recomenda-se que futuras pesquisas busquem ampliar a base amostral ou adaptar metodologias para melhor contemplar esse segmento empresarial.

REFERÊNCIAS BIBLIOGRÁFICAS

Agarwal, Vineet; Taffler, Richard. Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking & Finance*, v. 32, n. 8, p. 1541-1551, 2008.

Agarwal, Vineet., Bauer, Julian; Are hazard models superior to traditional bankruptcy prediction approaches? A comprehensive test. *Journal of Banking & Finance*, v. 40, p. 432-442, 2014. ISSN 0378-4266.

Akerlof, George A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84, no. 3 (1970): 488–500. https://doi.org/10.2307/1879431.

Allaire J, Chollet F (2024). keras: R Interface to 'Keras'. R package version 2.15.0, https://CRAN.R-project.org/package=keras.

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. The Journal of Finance.

Altman, E. I. (1984). The success of business failure prediction models: An international survey. Journal of Banking & Finance.

Altman, Edward I., 2005. "An emerging market credit scoring system for corporate bonds," Emerging Markets Review, Elsevier, vol. 6(4), pages 311-323, December.

Almeida, C. & Meres, B. (2008). Extracting Default Probabilities from Sovereign Bonds. Brazilian Review of Econometrics.

ABJ, Associação Brasileira de Jurimetria. Observatório da Insolvência. Projeto contínuo em parceria com o Núcleo de Estudos de Processos de Insolvência (NEPI) da PUC-SP. Início: 2017. Dados fornecidos sobre falências nos estados do RJ, SP e RS. Estudo disponível em: https://www.abj.org.br. Acesso em: 08 dez. 2024.

Beaver, William H. Financial ratios as predictors of failure. Journal of accounting research, v. 4, n. 3, p. 71-111, 1966.

Barboza, Flavio; kimura, Herbert; Altman, E.. Machine learning models and bankruptcy prediction. Expert Systems with Applications, v. 83, p. 405-417, 2017.

Blöchlinger, Andreas; leippold, Markus. Economic benefit of powerful credit scoring. Journal of Banking & Finance, v. 30, n. 3, p. 851-873, 2006.

Breiman, L. (2001). Random forests, Machine Learning 45: 5–32

Burgt, M. J**.** Calibrating low-default portfolios, using the cumulative accuracy profile. *Journal of Risk Model Validation*, v. 1, n. 4, p. 1-17, Winter 2007.

Canbas, S., Cabuk, A., & Kilic, S. B. (2005). Prediction of Commercial Bank Failure via Multivariate Statistical Analysis of Financial Structures: The Turkish Case. European Journal of Operational Research.

CNPJá API. Disponível em: https://cnpja.com/. Acesso em: 30 de agosto de 2024.

Cox, D. R. (1958). The regression analysis of binary sequences. Journal of the Royal Statistical Society: Series B (Methodological), 20(2), 215–242

Custódio, Cláudia; Ferreira, Miguel A.; Garcia-Appendini, Emilia. Indirect costs of financial distress. *Review of Finance*, v. 27, p. 2233-2270, 2023

Data science central. ROC Curve Explained in One Picture. Disponível em: https://www.datasciencecentral.com/roc-curve-explained-in-one-picture/. Acesso em: 20 nov. 2024.

DeMarzo, Peter; Berk, Jonathan; Finanças empresariais, essencial. Tradução de Christiane de Brito Andrei. 1. ed. Porto Alegre: Bookman, 2010. Tradução de: *Corporate Finance, the Core*. Pearson Education Inc., 2009

Dorogush, A.; Ershov, V.; Gulin, A. CatBoost: gradient boosting with categorical features support. 2018. Disponível em: https://arxiv.org/abs/1810.11363.

Elton, Edwin J.; Gruber, Martin J.; Blake, Christopher R. Survivorship bias and mutual fund performance. Review of Financial Studies, v. 9, n. 4, p. 1097-1120, 1996.

Fávero, Luiz Paulo; Belfiore, Patrícia. *Manual de Análise de Dados: Estatística e Modelagem Multivariada com Excel®, SPSS® e Stata®*. 1. ed. São Paulo: Elsevier, 2017.

Friedman, J., Hastie, T. and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent, Journal of Statistical - Software 33 (1): 1–22.

Frydman, H., Altman, E.I. and Kao, D. "Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress". *The Journal of Finance*. V. XI, n. 1 (March), p. 269-291, 1985.

Gneiting, T., Balabdaoui, F., & Raftery, A. E. (2007). Probabilistic forecasts, calibration and sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69* (2), 243–268 .

Hillegeist, S. A., Keating, E. K., Cram, D. P., & Lundstedt, K. G. (2004). Assessing the Probability of Bankruptcy. Review of Accounting Studies.

Jabeur, Sami Ben; gharib, Cheima; Mefteh-Wali, Salma; ben arfi, Wissal. CatBoost model and artificial intelligence techniques for corporate failure prediction. Technological Forecasting and Social Change, v. 166, 2021, 120658. ISSN 0040-1625.

James, G., D. Witten, T. Hastie, and R. Tibshirani. 2023.An introduction to statistical learning, 2nd ed. New York: Springer.

Jardin, P. du (2015). Bankruptcy Prediction Using Terminal Failure Processes. European Journal of Operational Research.

Jardin, P. du (2016). A Two-Stage Classification Technique for Bankruptcy Prediction. European Journal of Operational Research.

Jones, S. (2017). Corporate Bankruptcy Prediction: A High Dimensional Analysis. Journal of Business Finance & Accounting.

Jorda, Òscar; Knoll, Katharina; Kuvshinov, Dmitry; Schularick, Moritz; Taylor, Alan M. The rate of return on everything, 1870–2015. The Quarterly Journal of Economics, v. 134, n. 3, p. 1225-1298, 2019.

Jupetipe, Fernanda Karoliny Nascimento; Martins, Eliseu; Mário, Poueri do Carmo; Carvalho, Luiz Nelson Guedes de. Custos de falência no Brasil comparativamente aos estudos norte-americanos. *Revista Direito GV*, v. 13, n. 1, abr. 2017. DOI: 10.1590/2317-6172201702

Kumar, P. R., & Ravi, V. (2007). Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A review. European Journal of Operational Research.

Lohmann, Christian; Möllenhoff, Steffen; Ohliger, Thorsten. Nonlinear relationships in bankruptcy prediction and their effect on the profitability of bankruptcy prediction models. *Journal of Business Economics*, v. 93, p. 1661-1690, 2023.

Mair P, De Leeuw J, Hornik K (2023). _isotone: Active Set and Generalized PAVA for Isotone Optimization_. R package version 1.1-1, <https://CRAN.R-project.org/package=isotone>

Melo, L. L. de; Bortoluzzo, A. B.; Venezuela, M. K. Relevância do fluxo de caixa operacional para previsão de falência das empresas brasileiras abertas. Revista de Gestão dos Países de Língua Portuguesa, Rio de Janeiro, v. 21, n. 2, p. 67–89, 2022.

Mensah, Y. M. (1984). An Examination of the Stationarity of Multivariate Bankruptcy Prediction Models: A Methodological Study. Journal of Accounting Research.

Merton, R. C. On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance*, v. 29, n. 2, p. 449-470, 1974

Modigliani, Franco; Miller, Merton H. The cost of capital, corporation finance, and the theory of investment: reply. *The American Economic Review*, v. 49, n. 4, p. 655-669, set. 1959.

Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. Journal of Accounting Research.

Platt, J., et al., (1999). Probabilistic outputs for support vector machines and com-parisons to regularized likelihood methods. *Advances in Large Margin Classifiers, 10* (3), 61–74 .

Prokhorenkova, L.; Gusev, G.; vorobev, A.; Dorogush, A.; Gulin, A. CatBoost: unbiased boosting with categorical features. 2017. Disponível em: https://arxiv.org/abs/1706.09516

Rajan, U., Seru, A., & Vig, V. (2015). The Failure of Models That Predict Failure: Distance, Incentives, and Defaults. Journal of Financial Economics.

RFB, Receita Federal do Brasil, Redesim - Rede Nacional para a Simplificação do Registro e da Legalização de Empresas e Negócios. Disponível em: https://www.gov.br/empresas-e-negocios/pt-br/redesim. Acesso em: 30 de agosto de 2024.

Serasa, Indicadores Econômicos. Disponível em: https://www.serasaexperian.com.br/conteudos/indicadores-economicos/. Acesso em: 30 de novembro de 2024.

Serasa, Experian, Sala de Imprensa, Indicadores. Disponível em: https://www.serasaexperian.com.br/sala-de-imprensa/indicadores/mais-de-mil-pedidos-de-recuperacao-judicial-sao-registrados-no-semestre-mostra-serasa-experian/ . Acesso em: 30 de november de 2024.

Shi, Y., Ke, G., Soukhavong, D., Lamb, J., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T., Titov, N., & Cortes, D. (2024). *lightgbm: Light Gradient Boosting Machine*. R package version 4.5.0. Disponível em: https://CRAN.R-project.org/package=lightgbm.

Sigrist, F., & Leuenberger, N. (2022). Machine learning for corporate default risk: Multi-period prediction, frailty correlation, loan portfolios, and tail probabilities. Journal of Banking & Finance.

Tian, S., Yu, Y., & Guo, H. (2015). Variable selection and corporate bankruptcy forecasts. Journal of Banking & Finance, 52, 89–100

Valor Pro**.** Valor Econômico - Disponível em: https://valorpro.globo.com. Acesso em: agosto de 2024.

Wright, M. N., & Ziegler, A. (2017). ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. Journal of Statistical Software, 77(1), 1-17. https://doi.org/10.18637/jss.v077.i01

Zadrozny, B., & Elkan, C. (2001). Obtaining calibrated probability estimates from decision trees and naive Bayesian classifiers. In *Icml: vol. 1* (pp. 609–616). Citeseer.

Zeidan, R., & Onabolu, S. (2023). The Generalized Sustainability Credit Rating System. Journal of Sustainable Finance & Investment.

Zhang, G., Hu, M. Y., Patuwo, B. E., & Indro, D. C. (1997). Artificial Neural Networks in Bankruptcy Prediction: General Framework and Cross-Validation Analysis. European Journal of Operational Research.

Zuppini, M. S., Pinto, A. C., & Matsumoto, É. Y. (2021). Apreçamento de debêntures ilíquidas combinando técnicas de aprendizado não supervisionado e supervisionado. Revista Brasileira de Finanças.

APÊNDICES

Apêndice A - AUC para diferentes horizontes de previsão (em trimestres) e métodos.

Trimestre	Regressão Logística	Random Forest	CatBoost	LightGBM	Z-score (EMS)
1	0,9073	0,9053	0,9001	0,7743	0,8670
2	0,9081	0,9088	0,8885	0,7999	0,8598
3	0,9099	0,9084	0,8735	0,7685	0,8123
4	0,8961	0,9053	0,9044	0,8116	0,8154
5	0,8866	0,8847	0,8375	0,7351	0,7989
6	0,8745	0,8797	0,8810	0,6616	0,7867
7	0,8718	0,8746	0,8682	0,6204	0,7779
8	0,8588	0,8720	0,8491	0,6654	0,7677
9	0,8486	0,8718	0,7877	0,6085	0,7678
10	0,8544	0,8661	0,7768	0,6186	0,7613
11	0,8566	0,8588	0,7525	0,4774	0,7711
12	0,8280	0,8432	0,7287	0,5307	0,7345

Fonte: (AUTOR, 2024).

Apêndice B - Tabela dos hiperparâmetros

No modelo Random Forest, foi realizada uma análise inicial baseada no erro OOB (Out-of-Bag) para determinar o número ótimo de árvores. O teste OOB foi conduzido no primeiro horizonte de previsão, avaliando o desempenho do modelo para diferentes quantidades de árvores, variando de 100 a 10.000 em incrementos de 500. O número de árvores que resultou no menor erro OOB foi identificado como 3.600. Essa configuração foi mantida como padrão para os horizontes subsequentes.

Após a etapa OOB, foram identificadas as cinco variáveis mais relevantes para divisão dos nós no primeiro horizonte: CX_AT, AC_PC, PL_AT, PT_AT e RL_AT. Essas variáveis foram utilizadas para gerar combinações de dois e três elementos, totalizando 50 combinações diferentes. Cada combinação foi avaliada com base no AUC de validação, forçando o modelo a considerar essas variáveis no parâmetro always.split.variables. A combinação que apresentou o maior AUC foi CX_AT e AC_PC, e foi selecionada para os demais horizontes.

Para os demais modelos a escolha dos hiperparâmetros seguiu a regra mencionada no item 3.7, ou seja, o conjunto de hiperparâmetros que maximizasse o valor de AUC no conjunto de validação.

Todos os modelos desse trabalho foram testados em um notebook com processador 13th Gen Intel Core i7-1365U, 1.80 GHz, 32GB RAM, placa de vídeo Intel Iris Xe Graphics Family.

As tabelas a seguir detalham os hiperparâmetros ajustados para cada modelo, bem como as configurações específicas aplicadas durante o treinamento.

Hiperparâmetros testados para Modelo Regressão Logística

	Valores Testados
Família de distribuição	binomial
Método de regularização	alpha = 0 (Lasso: L1)
Lambda (Regularização)	Validação cruzada em cv.glmnet com lambda.min
Número de folds	nfolds = 100
Pesos de classes	1 / sum(y == 1) para default, 1 / sum(y == 0) para solventes
Padronização de variáveis	standardize = TRUE
Iteração por horizonte (H)	{1, 2, ..., 12}
Semente Aleatória (random_seed)	123
Tempo de processamento para 12 horizontes (minutos)	164

Fonte: (AUTOR, 2024).

Hiperparâmetros testados para Modelo Catboost	
	Valores Testados
Profundidade (depth)	{1, 4, 6, 8}
Taxa de Aprendizado (learning_rate)	{0.01, 0.1}
Regularização L2 (l2_leaf_reg)	{1, 5, 10}
Fracção de Atributos (rsm)	{0.1, 0.5, 1.0}
Contagem de Divisões (border_count)	{1, 64, 128, 254}
Iterações (iterations)	{100, 1000, 3000}
Função de Perda (loss_function)	Logloss
Semente Aleatória (random_seed)	123
Número de iterações para aguardar melhoria antes de parar (od_wait)	{20}
Tempo de processamento para 12 horizontes (minutos)	180

Fonte: (AUTOR, 2024).

Hiperparâmetros testados para LigthGBM	
	Valores Testados
NumLeaves	{31, 63, 127}
MaxDepth	{-1, 10}
Learning Rate	{0.1}
FeatureFraction	{1.0, 0.5}
BaggingFraction	{1.0, 0.5}
BaggingFreq	{0, 1, 5}
NumIterations	{100, 1000}
Lambda_L1	{0.1, 0.5}
Lambda_L2	{0.1, 0.5}
IsUnbalance	TRUE
EarlyStoppingRounds	{20}
Tempo de processamento para 12 horizontes (minutos)	64

Hiperparâmetros testados para Modelo Random Forest	
	Valores Testados
Número de Árvores	{100, 600, 1100, ..., 10000} (incrementos de 500) | 3600*
Profundidade Máxima (max.depth)	{5, 10, 15}
Número Mínimo de Amostras por Nó	1
Número de Variáveis para Divisão (mtry)	sqrt(p), onde p = número de variáveis (aproximado pelo código)
Frações de Amostras (sample.fraction)	{(0.1,0.3), c(0.1,0.5), c(0.1,0.7), c(0.2,0.3), c(0.2,0.5), c(0.2,0.7), c(0.5,0.3), c(0.5,0.5), c(0.5,0.7)}
Reamostragem (replace)	TRUE
Regra de Divisão (splitrule)	"extratrees"
Pesos das Classes (class.weights)	Ajustados dinamicamente como {classe 0 = 1, classe 1 = n0/n1}
Variáveis Fixas para Divisão	{"CX_AT_lag", "AC_PC_lag"}
Semente Aleatória (random_seed)	123
Tempo de processamento para 12 horizontes (minutos)	159

Fonte: (AUTOR, 2024).

Apêndice C – Distribuição da Amostra por Porte e Natureza Jurídica

Tabela 12: Distribuição da amostra por Porte e Natureza Jurídica.

Faturamento anual maior R$4,8 milhões = Média/Grande

Empresas Solventes (número de empresas)	**Micro e Pequenas**	**Grande e Média**	**Total**
Privadas com Fins Lucrativos (Aberta)	12	528	540
Privadas com Fins Lucrativos (Fechada)	365	5.437	5.802
Entidades Públicas e Mistas	8	171	179
Privadas s/ Fins Lucrativos/ Especiais	1	222	223
SubTotal Solventes	**386**	**6.358**	**6.744**

Empresas Default (número de empresas)	**Micro e Pequenas**	**Grande e Média**	**Total**
Privadas com Fins Lucrativos (Aberta)	3	67	70
Privadas com Fins Lucrativos (Fechada)	12	341	353
Entidades Públicas e Mistas	0	4	4
Privadas s/ Fins Lucrativos/ Especiais	0	9	9
SubTotal Solventes	**15**	**421**	**436**

Tabela 13: Distribuição da amostra por Porte e Setor.

Faturamento anual maior R$4,8 milhões = Média/Grande

Empresas Solventes (número de empresas)	Micro e Pequenas	Grande e Média	Total
Setor de comércio	18	678	696
Indústria	49	1.538	1.587
Setor de serviços	283	3.106	3389
Setor de energia	25	854	879
Agropecuária	11	182	193
SubTotal Solventes	**386**	**6.358**	**6.744**

Empresas Solventes (número de empresas)	Micro e Pequenas	Grande e Média	Total
Setor de comércio	0	38	38
Indústria	5	166	171
Setor de serviços	9	145	154
Setor de energia	0	62	62
Agropecuária	1	10	11
SubTotal Solventes	**15**	**421**	**436**

ANEXOS

Anexo A - Teste McNemar

O teste de McNemar é um método estatístico não paramétrico que pode ser utilizado para avaliar se há uma diferença significativa entre dois classificadores testados no mesmo conjunto de dados (Dietterich, 1998).

No contexto deste estudo, o teste foi aplicado para comparar o modelo Altman (EMS) e o modelo Random Forest. A análise concentra-se nos casos em que os dois modelos divergem em suas previsões, avaliando se o número de situações em que o modelo Random Forest acerta enquanto o modelo EMS erra difere significativamente do número de situações opostas, onde o modelo EMS acerta e o Random Forest erra.

As hipóteses do teste são formuladas da seguinte maneira:

- Hipótese nula (H_0): Não há diferença significativa entre as taxas de erro do modelo EMS e do modelo Random Forest. Qualquer discrepância entre as previsões dos modelos ocorre por acaso.
- Hipótese alternativa (H_a): Existe uma diferença significativa entre as taxas de erro do modelo EMS e do modelo Random Forest.

Para avaliação do teste os resultados das previsões dos dois modelos são organizados em uma tabela 2 x 2, conforme a tabela abaixo:

	Random Forest Correto	Random Forest Errado	Total
EMS Correto	a	b	$a + b$
SEM Errado	c	d	$c + d$
Total	$\boldsymbol{a + c}$	$\boldsymbol{b + d}$	$\boldsymbol{n}$

Fonte: (AUTOR, 2024), adaptado de Fávero et.al (2017, p.259)

Para testar a hipótese nula, a estatística do teste de McNemar é calculada usando a fórmula:

$$\chi^2 = \frac{(b - c)^2}{b + c}$$

onde b representa os casos em que o modelo Altman (EMS) erra e o Random Forest acerta, enquanto c representa os casos em que o Random Forest erra e o EMS acerta. Essa

estatística segue uma distribuição qui-quadrado com um grau de liberdade Fávero et.al (2017, p.259). Se o valor da estatística ultrapassar um limite crítico pré-definido, rejeitamos a hipótese nula em favor da alternativa, indicando que as taxas de erro dos dois modelos são significativamente diferentes.

Anexo B - Cálculo dos ativos ponderados pelo risco (Risk-Weighted Assets, RWA)

A fórmula para o cálculo dos Ativos Ponderados pelo Risco (Risk-Weighted Assets, RWA) foi retirada do documento *International Convergence of Capital Measurement and Capital Standards: A Revised Framework* (Basel Committee on Banking Supervision, 2004), disponível na seção dedicada às exposições corporativas, soberanas e bancárias (p. 71). Essa metodologia, também foi utilizada por Lohmann (2022), para avaliar os requisitos de capital de acordo com o framework Basel III. A fórmula incorpora parâmetros como Probabilidade de Default (PD), Perda Dada Default (LGD), Exposição no Momento do Default (EAD) e Maturidade Efetiva (M), ajustados por fatores de correlação e funções normais;

- Correlação (R):

$$R = 0.12 \cdot \left(\frac{1 - e^{-5 \ .PD}}{1 - e^{-5}}\right) + 0.24 \cdot \left(1 - \frac{1 - e^{-50.PD}}{1 - e^{-50}}\right)$$

- Ajuste de Maturidade (b)

$$b = (0.11852 - 0.05478 \cdot \ln(PD))^2$$

- Requisito de capital (K)

$$k = \left[LGD \cdot N\left((1 - R)^{-0.5} \cdot G(PD) + \sqrt{\frac{R}{1-R}} \cdot G(0.999)\right) - PD \cdot LGD\right] \cdot (1 - 1.5 \cdot b)^{-1} \cdot (1 + (M - 2.5) \cdot b)$$

- Ativos ponderados pelo risco (RWA)

$$RWA = k \cdot 12.5 \cdot EAD$$

Onde PD = probabilidade de default. LGD = perda dado o default, $N(.)$= função de densidade normal cumulativa, $G(.)$ = função inversa da densidade normal cumulativa, M = maturidade efetiva e EAD = exposição no momento do default.