



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS E CIÊNCIAS
ECONÔMICAS
CURSO DE CIÊNCIAS CONTÁBEIS

GABRIEL DE SOUSA MENDES

APLICAÇÃO DE REDES NEURAS NA PREVISÃO DE DIFICULDADES
FINANCEIRAS EM EMPRESAS LISTADAS NA B3

GOIÂNIA - GO
2025

Processo: 23070.059660/2025-35
Documento: 5841816



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS E CIÊNCIAS ECONÔMICAS

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR VERSÕES ELETRÔNICAS DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO NO REPOSITÓRIO INSTITUCIONAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio do Repositório Institucional (RI/UFG), regulamentado pela Resolução CEPEC no 1240/2014, sem ressarcimento dos direitos autorais, de acordo com a Lei no 9.610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou download, a título de divulgação da produção científica brasileira, a partir desta data.

O conteúdo dos Trabalhos de Conclusão dos Cursos de Graduação disponibilizado no RI/UFG é de responsabilidade exclusiva dos autores. Ao encaminhar(em) o produto final, o(s) autor(a)(es)(as) e o(a) orientador(a) firmam o compromisso de que o trabalho não contém nenhuma violação de quaisquer direitos autorais ou outro direito de terceiros.

1. Identificação do Trabalho de Conclusão de Curso de Graduação (TCCG)

Nome(s) completo(s) do(a)(s) autor(a)(es)(as): **GABRIEL DE SOUSA MENDES**

Título do trabalho: **APLICAÇÃO DE REDES NEURAIAS NA PREVISÃO DE DIFICULDADES FINANCEIRAS EM EMPRESAS LISTADAS NA B3**

2. Informações de acesso ao documento (este campo deve ser preenchido pelo orientador) Concorda com a liberação total do documento SIM NÃO¹

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. Após esse período, a possível disponibilização ocorrerá apenas mediante: a) consulta ao(à)(s) autor(a)(es)(as) e ao(à) orientador(a); b) novo Termo de Ciência e de Autorização (TECA) assinado e inserido no arquivo do TCCG. O documento não será disponibilizado durante o período de embargo.

Casos de embargo:

- Solicitação de registro de patente;
- Submissão de artigo em revista científica;
- Publicação como capítulo de livro.

Obs.: Este termo deve ser assinado no SEI pelo orientador e pelo autor.



Documento assinado eletronicamente por **Carlos Henrique Silva Do Carmo, Professor do Magistério Superior**, em 08/12/2025, às 10:20, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Gabriel De Sousa Mendes, Discente**, em 08/12/2025, às 10:28, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5841816** e o código CRC **6941008E**.

GABRIEL DE SOUSA MENDES

APLICAÇÃO DE REDES NEURAIS NA PREVISÃO DE DIFICULDADES
FINANCEIRAS EM EMPRESAS LISTADAS NA B3

Trabalho de Conclusão de Curso
apresentado ao curso de Ciências
Contábeis da Universidade Federal
de Goiás - UFG, como requisito
parcial para a obtenção do título de
Bacharel em Ciências Contábeis.

Orientador: Prof. Dr. Carlos
Henrique Silva do Carmo

GOIÂNIA - GO
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Mendes, Gabriel de Sousa
APLICAÇÃO DE REDES NEURAIIS NA PREVISÃO DE
DIFICULDADES FINANCEIRAS EM EMPRESAS LISTADAS NA B3
[manuscrito] / Gabriel de Sousa Mendes. - 2025.
44 f.: il.

Orientador: Prof. Dr. Carlos Henrique Silva do Carmo.
Trabalho de Conclusão de Curso (Graduação) - Universidade
Federal de Goiás, Faculdade de Administração, Ciências Contábeis
e Ciências Econômicas (FACE), Ciências Contábeis, Goiânia, 2025.
Bibliografia.

Inclui gráfico, tabelas, lista de figuras, lista de tabelas.

1. Previsão de Insolvência. 2. Financial distress. 3. Redes Neurais
Artificiais. 4. Modelos Preditivos. 5. Engenharia de Feature. I. Carmo,
Carlos Henrique Silva do, orient. II. Título.

CDU 005



UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS E CIÊNCIAS ECONÔMICAS

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO II DE CIÊNCIAS CONTÁBEIS (TCC II)

Ao(s) vinte e quatro dia(s) do mês de novembro do ano de dois mil e vinte e cinco iniciou-se a sessão pública de defesa do Trabalho de Conclusão de Curso II (TCC II) intitulado “**APLICAÇÃO DE REDES NEURAIS NA PREVISÃO DE DIFICULDADES FINANCEIRAS EM EMPRESAS LISTADAS NA B3**”, de autoria do discente **GABRIEL DE SOUSA MENDES**, matrícula **201710163** do curso de Ciências Contábeis, da Faculdade de Administração, Ciências Contábeis e Ciências Econômicas da UFG. Os trabalhos foram instalados pelo presidente da banca **CARLOS HENRIQUE SILVA DO CARMO (FACE/UFG)** com a participação dos demais membros da Banca Examinadora: **ELMO DIAS DA SILVEIRA (CEPECONF/UFG)** e **ERCÍLIO ZANOLLA (FACE/UFG)**.

Após exposição de quinze minutos, o discente foi arguido oralmente pelos membros da Banca Examinadora. Nesta arguição a Banca buscou aferir a suficiência de conhecimento e a capacidade de sistematização do tema desenvolvido pelo discente em seu TCC II. Após realização dos comentários de cada um dos professores examinadores, a Banca reuniu-se reservadamente e atribuiu a nota final **10,0 (Dez)**, tendo sido o TCC II considerado **Aprovado**.

Proclamados os resultados, os trabalhos foram encerrados e, para constar, lavrou-se a presente ata que segue assinada pelos membros da Banca Examinadora.



Documento assinado eletronicamente por **Carlos Henrique Silva Do Carmo, Professor do Magistério Superior**, em 24/11/2025, às 20:29, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ercilio Zanolla, Professor do Magistério Superior**, em 24/11/2025, às 20:56, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Elmo Dias da Silveira, Usuário Externo**, em 25/11/2025, às 09:03, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5808664** e o código CRC **701E1875**.

RESUMO

A capacidade de antecipar dificuldades financeiras é um pilar da gestão de risco e análise de investimentos, sendo que os modelos estatísticos tradicionais apresentam limitações significativas em ambientes econômicos voláteis. Este trabalho teve como objetivo central avaliar a aplicação e a eficácia de Redes Neurais Artificiais (RNA), mais especificamente um *Perceptron* de Múltiplas Camadas (MLP), na previsão de *financial distress* em empresas listadas na B3, com um horizonte preditivo de um ano. A metodologia foi baseada em um *pipeline* rigoroso, partindo de 9.025 observações trimestrais (2010-2024), que incluiu a construção de uma variável-alvo temporalmente defasada e uma engenharia de *features* híbrida, combinando indicadores financeiros, macroeconômicos e de controle setorial. A performance do MLP foi comparada diretamente a um benchmark de Regressão Logística e aos modelos tradicionais internacionais como o *z-score* de Altman e nacionais como os de Elizabetsky e de Kanitz e através de validação cruzada temporal e um teste final em dados recentes. Os resultados indicam que a engenharia de *features* híbrida é o principal motor de performance, elevando o modelo Logit a um AUC de 0.9070. Contudo, a arquitetura não-linear do MLP demonstrou uma vantagem estatisticamente significativa no teste final, atingindo um AUC de 0.9229. A análise de interpretabilidade confirmou que ambos os modelos concordam sobre as *features* mais importantes, validando o conjunto híbrido e sugerindo que a vantagem do MLP reside na sua capacidade de capturar interações não-lineares complexas. Conclui-se que o MLP é a ferramenta de melhor desempenho, mas seu sucesso é fundamentalmente dependente de uma engenharia de *features* que inclua o contexto macroeconômico.

Palavras-chave: Previsão de Insolvência. *Financial distress*. Redes Neurais Artificiais. Modelos Preditivos. Engenharia de *Features*. B3.

ABSTRACT

The ability to anticipate financial distress is a pillar of risk management and investment analysis, as traditional statistical models show significant limitations in volatile economic environments. This study's central objective was to evaluate the application and effectiveness of Artificial Neural Networks (ANN), more specifically a Multilayer Perceptron (MLP), in predicting financial distress for companies listed on the B3, with a one-year predictive horizon. The methodology was based on a rigorous pipeline that started with 9,025 quarterly observations (2010-2024), including the construction of a time-lagged target variable and a hybrid feature engineering approach that combined financial, macroeconomic, and sectoral control indicators. The MLP's performance was compared directly to a Logistic Regression benchmark and to international historical models, such as Altman's z-score, and national models, such as those of Elizabetsky and Kanitz, through temporal cross-validation and a final test on recent data. The results indicate that hybrid feature engineering is the primary driver of performance, elevating the Logit model's AUC to 0.9070. However, the non-linear architecture of the MLP demonstrated a statistically significant advantage in the final test, achieving an AUC of 0.9229. The interpretability analysis confirmed that both models agree on the most important features, validating the hybrid set and suggesting that the MLP's advantage lies in its ability to capture complex nonlinear interactions. It is concluded that the MLP is the superior-performing tool, but its success is fundamentally dependent on feature engineering that incorporates macroeconomic context.

Keywords: Insolvency Forecasting. Financial distress. Artificial Neural Networks. Predictive Models. Feature Engineering. B3.

LISTA DE TABELAS

Tabela 1: Execução do funil de filtragem da amostra	19
Tabela 2: Distribuição da base por setor.....	20
Tabela 3: Distribuição da base por ano.....	21
Tabela 4: Métricas da validação cruzada K-Fold	34
Tabela 5: Métricas da validação final entre os dois modelos	36
Tabela 6: Performance por setor.....	37

LISTA DE QUADROS

Quadro 1: Modelos preditivos de insolvência tradicionais	14
Quadro 2: Componentes dos modelos tradicionais	15
Quadro 3: Variáveis independentes	24
Quadro 4: Arquitetura e hiper parâmetros do modelo	29
Quadro 5: Estratégia de avaliação comparativa dos modelos	30
Quadro 6: Consistência das features.....	35

LISTA DE FIGURAS

Figura 1: Comparativo entre importância das variáveis para cada modelo.....	38
Figura 2: Importância das features para o modelo MLP	40
Figura 3: Importância das features para o modelo LOGIT	40

SUMÁRIO

1 INTRODUÇÃO.....	10
2 REVISÃO DE LITERATURA.....	12
2.1 Emergência da preocupação com dificuldades financeiras	12
2.2 Primeiras propostas: Modelos estatísticos tradicionais	13
2.3 Avanço tecnológico e redes neurais artificiais.....	16
2.4 Caminhos atuais e direções futuras	17
3 METODOLOGIA.....	18
3.1 Análise exploratória e amostragem dos dados.....	18
3.2 Construção da variável dependente financial distress	21
3.3 Engenharia de features e variáveis de controle	22
3.4 Pipeline de modelagem e validação cruzada	27
3.5 Interpretabilidade do modelo.....	31
3.6 Métricas de avaliação de performance	31
4 RESULTADOS E DISCUSSÃO	32
4.1 Performance comparativa dos modelos	32
4.2 Análise de interpretabilidade do modelo SHAP	38
5 CONSIDERAÇÕES FINAIS	40
REFERÊNCIAS	42

1 INTRODUÇÃO

A capacidade de antecipar problemas financeiros nas empresas é de grande relevância para investidores, gestores e instituições financeiras, especialmente em economias em desenvolvimento, como é o caso do Brasil. Problemas financeiros não previstos adequadamente podem levar empresas à insolvência financeira, gerando impactos negativos em toda cadeia produtiva e afetando significativamente a confiança dos investidores (REZENDE et al., 2017).

A insolvência financeira não deve ser confundida com a falência, embora sejam frequentemente abordadas de maneira similar. Enquanto a falência é um estágio jurídico e final, regulado, assim como a recuperação judicial, pela **Lei de Recuperação de Empresas e Falência** (Lei nº 11.101/2005, atualizada pela Lei nº 14.112/2020), a insolvência financeira caracteriza-se por dificuldades persistentes em honrar compromissos financeiros antes mesmo de atingir a falência formal (Oshiro, 2016). Para fins deste trabalho, os termos "dificuldade financeira" e "*financial distress*" serão utilizados como sinônimos operacionais deste estado de insolvência financeira. Identificar precocemente tais dificuldades financeiras permite às organizações e seus stakeholders a adoção de medidas preventivas, evitando prejuízos maiores (REZENDE et al., 2017; BARBOZA; NOBRE, 2023).

Tradicionalmente, a literatura focou em duas eras de modelos estatísticos. A primeira, dos scores clássicos, produziu modelos como os de Altman (1968), Zmijewski (1984), e as adaptações brasileiras de Kanitz (1978), Matias (1978) e Elizabetsky (1976). A segunda era introduziu o padrão estatístico moderno, notadamente a Regressão Logística (Logit), popularizada por Ohlson (1980), que se tornou o benchmark linear dominante. Todos esses modelos baseiam-se na análise de indicadores extraídos das demonstrações contábeis (SOARES; MARIN; SANTOS, 2021). Embora historicamente relevantes, essas abordagens estatísticas apresentam limitações importantes quanto à sua capacidade de generalização e de adaptação a mudanças estruturais no ambiente econômico.

Uma das críticas centrais aos modelos tradicionais está na sua rigidez técnica. Esses métodos costumam assumir relações lineares entre variáveis e são fortemente dependentes da estabilidade dos coeficientes ao longo do tempo, o que compromete sua performance em contextos voláteis ou com mudanças significativas nas práticas contábeis e nos padrões de mercado (REZENDE et al., 2017). Além disso, modelos como o de

Altman não consideram interações complexas entre variáveis, tampouco incorporam dados de natureza não contábil, como indicadores macroeconômicos ou qualitativos (BARBOZA; NOBRE, 2023).

A literatura mais recente destaca a necessidade de incorporar variáveis de contexto econômico e institucional na modelagem do risco de insolvência. Rezende et al. (2017) evidenciam que fatores macroeconômicos e setoriais influenciam significativamente a saúde financeira das empresas. De forma complementar, Oshiro (2016) ressalta que fatores estruturais internos, como a concentração de decisões e baixa independência gerencial, também podem aumentar a propensão ao *financial distress*.

Diante dessas limitações, surgem como alternativa os modelos baseados em inteligência artificial, notadamente as redes neurais artificiais (RNAs). Este estudo foca na arquitetura mais utilizada e fundamental das RNAs, o *Perceptron* de Múltiplas Camadas (MLP). Essas técnicas têm a capacidade de aprender padrões a partir de grandes volumes de dados tradicionais, identificando relações não lineares e interações complexas entre variáveis financeiras, econômicas e de controle (LUNARDI; BILK; RODRIGUES JUNIOR, 2020; ELHOSENY et al., 2022). Tais características tornam as RNAs mais adaptáveis a diferentes contextos econômicos e potencialmente mais precisas na previsão de insolvência.

Além disso, esses modelos permitem maior flexibilidade na inclusão de variáveis, mesmo com distribuições não normais ou ausência de independência entre elas (LUNARDI; BILK; RODRIGUES JUNIOR, 2020). Entretanto, apesar das vantagens preditivas, a literatura também aponta desafios relacionados à interpretabilidade dos modelos baseados em redes neurais. A dificuldade em compreender os critérios de decisão utilizados pelas RNAs, frequentemente descrita como efeito “caixa-preta”, pode limitar sua aceitação em ambientes regulatórios e corporativos mais conservadores (Rudin, 2019).

Outro aspecto crítico identificado por Horta, Alves e Carvalho (2014) diz respeito à etapa de seleção de atributos. A escolha inadequada de variáveis pode comprometer tanto modelos estatísticos quanto modelos baseados em aprendizado de máquina. A ausência de um processo sistemático e fundamentado de seleção de indicadores reduz a capacidade preditiva dos modelos e dificulta sua replicação, o que reforça a importância de abordagens metodológicas rigorosas.

Diante das limitações dos modelos clássicos e da necessidade de uma abordagem mais robusta, que inclua fatores específicos do contexto brasileiro e indicadores macroeconômicos, este trabalho busca responder à seguinte pergunta de pesquisa: "Qual é o desempenho das redes neurais artificiais (não-lineares) em comparação ao benchmark linear moderno (Regressão Logística) na previsão de dificuldades financeiras em empresas listadas na B3?". Especificamente, objetiva-se avaliar se as RNAs oferecem uma precisão superior ao modelo Logit, quando ambos são treinados com o mesmo conjunto híbrido de indicadores financeiros, macroeconômicos e de controle setorial. Com isso, espera-se contribuir para a ampliação e o refinamento dos instrumentos de análise financeira no mercado brasileiro, oferecendo uma abordagem robusta e multidimensional para a previsão da insolvência financeira. A relevância destes achados reside na capacidade de fornecer subsídios tempestivos para que gestores e credores atuem preventivamente, alinhando-se ao princípio da preservação da empresa que rege a Lei de Recuperação de Empresas e Falência.

2 REVISÃO DE LITERATURA

2.1 Emergência da preocupação com dificuldades financeiras

A preocupação com a previsão de dificuldades financeiras em empresas, conhecida como *financial distress*, tem raízes históricas profundas. A Grande Depressão de 1929, por exemplo, expôs a fragilidade do sistema financeiro e a falta de ferramentas eficazes para avaliar a saúde das empresas e prever a iminência de falências. Como apontam Bellovary, Giacomino e Akers (2007), este evento impulsionou as primeiras pesquisas acadêmicas sobre o tema, que, na década de 1930, focaram no uso da análise de índices para identificar padrões que pudessem sinalizar o risco de insolvência. A partir de então, a capacidade de antecipar problemas financeiros consolidou-se como um dos principais temas em finanças corporativas, dada sua importância para a gestão de risco, formulação de estratégias e estabilidade do mercado.

O *financial distress* é conceituado como uma condição anterior à falência, caracterizada pela dificuldade recorrente de honrar compromissos financeiros. Segundo Platt e Platt (apud Oshiro, 2016), trata-se de um estágio de declínio operacional que antecede a insolvência formal, impactando negativamente o valor de mercado da empresa e dificultando o acesso a crédito. A identificação precoce dessas dificuldades torna-se uma vantagem competitiva para empresas e investidores. Modelos que conseguem capturar os

primeiros sinais de deterioração financeira permitem ajustes estratégicos e melhoria na alocação de recursos, além de reduzir a exposição à inadimplência.

Lunardi, Bilk e Rodrigues Junior (2020) apontam que o distanciamento entre a situação financeira atual da empresa e sua capacidade de liquidez gera desconfiança entre *stakeholders* e reduz a eficiência da gestão organizacional. Nesses contextos, a atuação preventiva com base em indicadores é essencial para mitigar perdas e evitar reestruturações forçadas.

2.2 Primeiras propostas: Modelos estatísticos tradicionais

Desde os anos 1960, diversos modelos estatísticos foram propostos para prever dificuldades financeiras. O modelo *Z-score*, originalmente proposto por Altman (1968), é citado por Soares et al. (2021) como pioneiro na utilização de análise discriminante multivariada para combinar variáveis contábeis, como liquidez, lucratividade e alavancagem, em uma fórmula linear. Posteriormente, Ohlson (1980) introduziu um modelo baseado em regressão logística, utilizando variáveis financeiras e categóricas para estimar a probabilidade de falência, aplicável a empresas de capital aberto em geral. Zmijewski (1984) propôs um modelo alternativo, baseado em regressão Probit, focado em três variáveis principais: rentabilidade, alavancagem e liquidez corrente, aplicado a em empresas não financeiras. Esses modelos tornaram-se amplamente utilizados devido à sua simplicidade e facilidade de implementação, além de fornecerem bons níveis de acurácia quando aplicados a contextos estáveis e bem definidos.

No Brasil, destacam-se os modelos desenvolvidos por Elizabetsky (1976), Matias (1978) e Kanitz (1978), que adaptaram técnicas similares ao contexto nacional. O modelo de Elizabetsky foi desenvolvido para auxiliar instituições financeiras na concessão de crédito, com base em índices como margem líquida, liquidez e estrutura patrimonial. Já o modelo de Matias foi direcionado ao setor industrial e busca facilitar a análise de crédito com um conjunto reduzido de indicadores financeiros, priorizando a aplicabilidade prática no mercado brasileiro. Kanitz propôs um modelo multivariado que utiliza índices contábeis de rentabilidade e liquidez para gerar um fator de insolvência, com aplicação voltada à análise de solvência empresarial no setor produtivo em geral. Posteriormente, outros autores ampliaram esse repertório com propostas mais atualizadas. Silva (1982) desenvolveu dois modelos distintos voltados aos setores comercial e industrial, com base em múltiplas variáveis contábeis. Sanvicente e Minardi (1998) introduziram ajustes metodológicos ao contexto brasileiro, utilizando técnicas de análise discriminante para

refinar os critérios de classificação. Além disso, Altman, Baidya e Dias (1979) adaptaram o modelo *Z-score* original à realidade das empresas brasileiras, apresentando variantes com pequenas alterações nas variáveis utilizadas, sem comprometer a comparabilidade com o modelo internacional.

Apesar dessas contribuições, os modelos tradicionais enfrentam limitações consideráveis. Barbosa e Nobre (2023) destacam que esses modelos apresentam baixo poder preditivo diante de relações complexas e não lineares, comuns em ambientes econômicos dinâmicos. Rezende et al. (2017) apontam que a utilização de variáveis exclusivamente contábeis limita a capacidade dos modelos de se adaptarem a diferentes contextos econômicos e setoriais. Essa dependência de dados tradicionais, muitas vezes defasados ou sujeitos a manipulação, compromete a precisão em ambientes instáveis ou de transformação rápida. Outro ponto de destaque é que esses modelos clássicos geralmente não incorporam variáveis qualitativas, como indicadores de governança ou fatores macroeconômicos. Com isso, reduzem sua aplicabilidade em mercados emergentes como o brasileiro, que demandam maior sensibilidade a contextos externos e institucionais (BARBOSA; NOBRE, 2023).

A ausência de variáveis macroeconômicas e a rigidez técnica desses modelos, somadas à maior disponibilidade de dados e poder computacional, justificaram a busca por abordagens mais flexíveis e adaptativas, como as redes neurais artificiais, que serão discutidas na seção seguinte.

Para consolidar a revisão dos *benchmarks* estatísticos, o **Quadro 1** resume as equações lineares, focos originais e limiares de risco dos principais modelos tradicionais adaptados ao contexto brasileiro, conforme utilizados neste trabalho. O **Quadro 2** detalha os componentes (variáveis de entrada) de cada modelo, explicitando as fórmulas contábeis utilizadas.

Quadro 1: Modelos preditivos de insolvência tradicionais

Modelo	Foco	Fórmula	Limiar
ALTMAN (1968)	EUA Indústria	$ALT = 1,2A1 + 1,4A2 + 3,3A3 + 0,6A4 + 1,0A5$	< 1.81
ELIZABETSKY (1976)	BRASIL Indústria	$ELI = 1,93E1 - 0,20E2 + 1,02E3 + 1,33E4 - 1,12E5$	≤ 0
KANITZ (1978)	BRASIL Multissetorial	$KAN = 0,05K1 + 1,65K2 + 3,55K3 - 1,06K4 - 0,33K5$	< 0
MATIAS (1978)	BRASIL Multissetorial	$MAT = 23,79M1 - 8,26M2 - 8,86M3 - 0,76M4 + 1,53M5 + 9,91M6$	≤ 0
ZMIJEWSKY (1984)	EUA Multissetorial	$ZMI = -4,336 + 4,513Z1 + 5,679Z2 - 0,004Z3$	> 0

Fonte: Elaborado pelo autor.

Os componentes de cada fórmula (A1, E1, K1, etc.) são indicadores financeiros clássicos. Conforme detalhado na Seção 3.3, este trabalho "desmonta" esses modelos e utiliza seus componentes individuais como *features* de entrada candidatas tanto para o modelo MLP quanto para a Regressão Logística. Adicionalmente, os scores finais (ALT, KAN, etc.) serão utilizados como *benchmarks* de comparação histórica, enquanto o modelo Logit servirá como o *benchmark* de arquitetura. O **Quadro 2** detalha estes componentes.

Quadro 2: Componentes dos modelos tradicionais

Modelo	Variável	Descrição	Fórmula
ALTMAN	A1	Capital de giro / Ativo total	$\frac{\text{Ativo circulante} - \text{Passivo circulante}}{\text{Ativo total}}$
	A2	Lucros retidos / Ativo total	$\frac{\text{Lucros Retidos}}{\text{Ativo total}}$
	A3	EBIT / Ativo total	$\frac{\text{EBIT}}{\text{Ativo total}}$
	A4	Valor de mercado / Passivo exigível	$\frac{\text{Valor ação} * \text{qtd ações}}{\text{Passivo exigível}}$
	A5	Giro do ativo	$\frac{\text{Receita líquida}}{\text{Ativo total}}$
ELIZABETSKY	E1	Margem líquida	$\frac{\text{Resultado líquido}}{\text{Receita líquida}}$
	E2	Caixa e equivalentes / Ativo permanente	$\frac{\text{Caixa e equivalentes}}{\text{Ativo permanente}}$
	E3	Contas a receber / Ativo total	$\frac{\text{Contas a receber}}{\text{Ativo total}}$
	E4	Estoques / Ativo total	$\frac{\text{Estoques}}{\text{Ativo total}}$
	E5	Passivo circulante / Ativo	$\frac{\text{Passivo circulante}}{\text{Ativo total}}$
KANITZ	K1	ROE	$\frac{\text{Resultado líquido}}{\text{Patrimônio líquido}}$
	K2	Liquidez geral	$\frac{\text{Ativo circulante} + \text{Realizável a LP}}{\text{Passivo exigível}}$
	K3	Liquidez seca	$\frac{\text{Ativo circulante} - \text{Estoques} - \text{Desp Antecip}}{\text{Passivo circulante}}$
	K4	Liquidez corrente	$\frac{\text{Ativo circulante}}{\text{Passivo circulante}}$
	K5	Grau de endividamento	$\frac{\text{Passivo exigível}}{\text{Patrimônio líquido}}$

MATIAS	M1	Patrimônio líquido / Ativo total	$\frac{\text{Patrimônio líquido}}{\text{Ativo total}}$
	M2	Dívida total / Ativo circulante	$\frac{\text{Dívida total}}{\text{Ativo circulante}}$
	M3	Fornecedores / Ativo total	$\frac{\text{Fornecedores}}{\text{Ativo total}}$
	M4	Liquidez corrente	$\frac{\text{Ativo circulante}}{\text{Passivo circulante}}$
	M5	EBIT / Resultado bruto	$\frac{\text{EBIT}}{\text{Resultado bruto}}$
	M6	Caixa e equivalentes / Ativo total	$\frac{\text{Caixa e equivalentes}}{\text{Ativo total}}$
ZMIJEWSKY	Z1	ROA	$\frac{\text{Resultado líquido}}{\text{Ativo total}}$
	Z2	Passivo exigível / Ativo total	$\frac{\text{Passivo exigível}}{\text{Ativo total}}$
	Z3	Liquidez corrente	$\frac{\text{Ativo circulante}}{\text{Passivo circulante}}$

Fonte: Elaborado pelo autor.

2.3 Avanço tecnológico e redes neurais artificiais

A Crise Financeira Global de 2008 expôs as limitações dos modelos estatísticos em contextos de alta complexidade, impulsionando a busca por abordagens mais sofisticadas. Com os avanços da tecnologia e a maior disponibilidade de dados, as redes neurais artificiais (RNAs) surgiram como uma alternativa capaz de superar as limitações dos modelos tradicionais. Elhoseny et al. (2022) destacam que as RNAs conseguem modelar relações não lineares entre variáveis, tornando-se altamente eficazes na previsão de eventos financeiros adversos. Lunardi, Bilk e Rodrigues Junior (2020) reforçam que essas técnicas aprendem padrões a partir de grandes volumes de dados, inclusive com variáveis interdependentes e sem pressupostos estatísticos rígidos.

Estudos recentes têm explorado o uso de modelos de aprendizado profundo e técnicas *ensemble* como alternativas aos modelos estatísticos convencionais. Irvan (2024), em sua análise comparativa na bolsa da Indonésia, demonstrou que os modelos baseados em redes neurais superam as abordagens tradicionais em termos de acurácia, especialmente na classificação de empresas em zonas de risco intermediário. Resultados semelhantes foram obtidos por Nguyen, Nguyen e Lieu (2023), que aplicaram redes neurais em empresas de capital aberto no Vietnã e comprovaram o desempenho superior das abordagens de *machine learning*. Em um estudo sobre empresas do setor de serviços na Índia, Sethi e Mahadik (2025) concluíram que o modelo baseado em RNA apresentou os melhores índices de desempenho. Na América Latina, Barboza e Altman (2024)

também demonstraram a superioridade de modelos como o *Random Forest* na previsão de insolvência.

Entretanto, apesar das vantagens preditivas, a literatura também aponta desafios relacionados à interpretabilidade dos modelos baseados em redes neurais. A dificuldade em compreender os critérios de decisão usados pelas RNAs, frequentemente descrita como efeito “caixa-preta”, pode limitar sua aceitação em ambientes regulatórios e corporativos mais conservadores (RUDIN, 2019). Essa limitação reside na dimensão da transparência e da auditabilidade: Enquanto modelos lineares oferecem coeficientes claros, as RNAs operam através de milhares de conexões não-lineares difíceis de rastrear. Para mitigar esse *trade-off* entre performance e explicabilidade, emergiu o campo de estudo da Inteligência Artificial Explicável (XAI). Nesse sentido, Tran et al. (2022) aplicaram técnicas de interpretação para avaliar a contribuição de cada variável nos modelos de previsão de distress, buscando compatibilizar performance com transparência nos resultados. Outro ponto sensível é a seleção de variáveis. Horta, Alves e Carvalho (2014) destacam que a ausência de um critério rigoroso para a escolha de atributos pode comprometer a capacidade preditiva, reforçando a necessidade de uma etapa cuidadosa de preparação e filtragem dos dados.

Além dos dados financeiros, estudos têm ampliado o escopo das variáveis utilizadas, incorporando fatores macroeconômicos. Rahayu, Kusuma e Arifin (2025) demonstraram que PIB, taxa de juros, inflação e câmbio afetam significativamente a ocorrência de *distress* em empresas listadas no sudeste asiático. A integração dessas variáveis em modelos de *machine learning* tem contribuído para aumentar a capacidade preditiva dos algoritmos em contextos voláteis e regionais.

Essas evidências reforçam que, embora os modelos baseados em inteligência artificial apresentem maior poder preditivo e capacidade de adaptação, sua utilização exige atenção especial à interpretabilidade, qualidade dos dados e coerência metodológica, aspectos centrais quando se considera sua aplicação prática no ambiente empresarial brasileiro.

2.4 Caminhos atuais e direções futuras

A evolução metodológica na previsão de dificuldades financeiras demonstra um movimento claro: dos modelos estatísticos lineares para abordagens baseadas em aprendizado de máquina. A literatura mais recente aponta que a efetividade desses

modelos depende não apenas da técnica empregada, mas também da qualidade dos dados, da seleção criteriosa de variáveis e da incorporação de fatores contextuais.

Nesse sentido, a discussão sobre modelos comparativos é fortalecida por autores como Nguyen, Nguyen e Lieu (2023), que exploram o desempenho relativo de abordagens estatísticas e de *machine learning* em economias de transição. Complementarmente, Tran et al. (2022) destacam a importância da interpretabilidade e da explicação dos resultados em modelos baseados em aprendizado de máquina.

De forma integrada, estudos como os de Barbosa e Nobre (2023) sustentam que a análise preditiva deve considerar os aspectos contábeis, institucionais e econômicos, principalmente em mercados com altos níveis de instabilidade, como o brasileiro. Além disso, Soares et al. (2021) reforçam a importância de modelos sensíveis ao contexto nacional e às diferentes características das empresas analisadas. A incorporação de fatores contextuais, como variáveis macroeconômicas e de governança, é crucial para uma análise mais precisa (REZENDE et al., 2017; OSHIRO, 2016).

Com isso, o desafio atual é desenvolver e aplicar modelos que combinem desempenho preditivo elevado com interpretabilidade e adequação ao ambiente institucional e econômico brasileiro. A presente pesquisa insere-se neste contexto, buscando testar a efetividade das redes neurais na previsão de dificuldades financeiras em empresas da B3, considerando indicadores financeiros e macroeconômicos.

3 METODOLOGIA

3.1 Análise exploratória e amostragem dos dados

Este estudo adotou uma abordagem quantitativa e preditiva. A constituição da amostra de dados é uma etapa metodológica crítica, pois define o universo no qual os modelos preditivos serão treinados e validados.

A fonte de dados primária foi a plataforma Economática, da qual foram extraídos dados cadastrais, de balanço patrimonial, de demonstração de resultados e de mercado. A população inicial analisada foi composta por empresas listadas na B3, totalizando 52.140 observações (empresa-trimestre) de 869 empresas.

O recorte temporal do estudo foi definido de 2010 a 2024. A escolha por 2010 como marco inicial se justifica pela adoção obrigatória das Normas Internacionais de Contabilidade (IFRS) no país, que consolidaram padrões contábeis e asseguraram maior comparabilidade entre as demonstrações financeiras, viabilizando o uso de dados desse período com confiabilidade. Foram utilizados dados trimestrais devido à sua maior

granularidade, o que permitiu observar com maior precisão os sinais de deterioração financeira, escolha esta compatível com estudos como Rezende et al. (2017) e Lunardi et al. (2020).

A escolha do recorte temporal também buscou contemplar um ciclo econômico completo, iniciando-se com a crise econômica de 2014–2016, seguida por uma recuperação gradual até 2019, o choque exógeno provocado pela pandemia da Covid-19 entre 2020 e 2021 e, mais recentemente, um novo ciclo de aperto monetário entre 2022 e 2023. Essa diversidade de contextos econômicos permitiu que o modelo fosse treinado em diferentes condições de estresse e normalidade.

Conforme a análise exploratória sistemática, foi aplicado um funil de filtragem sequencial para garantir a qualidade, comparabilidade e robustez da amostra. Este processo, alinhado às recomendações de Oshiro (2016) e Horta, Alves e Carvalho (2014), substitui regras de corte arbitrárias (como um percentual fixo de dados ausentes) por uma limpeza empírica baseada na validade de cada observação e empresa.

A Tabela 1 resume o impacto de cada critério de filtragem na constituição da amostra final.

Tabela 1: Execução do funil de filtragem da amostra

Etapa	Critério	Empresas removidas	Empresas restantes	Observações removidas	Observações restantes
-	Base inicial	-	869	-	52.140
1	Setor	478	391	28.680	23.460
2	Variável alvo	10	381	7.551	15.909
3	Volume mínimo	48	333	554	15.355
4	Viés de seleção	143	190	6.330	9.025

Fonte: Dados da pesquisa.

O processo de filtragem (Tabela 1) resultou em uma amostra final robusta e focada na previsão, composta por 9.025 observações de 190 empresas operacionais. A seguir, cada etapa do funil é detalhada.

A Etapa 1 (Filtro Setorial) removeu 478 entidades. Optou-se pela exclusão do setor financeiro, como bancos, seguradoras e corretoras, em linha com os procedimentos de Oshiro (2016) e Lunardi, Bilk e Rodrigues Junior (2020). A literatura especializada aponta que as instituições financeiras possuem características contábeis, operacionais e regulatórias específicas, que requerem modelagens próprias para avaliação de risco, não sendo comparáveis às demais empresas. Adicionalmente, esta etapa removeu empresas classificadas como "-" (sem classificação) ou "OUTROS", que, após análise exploratória,

foram identificadas como holdings não-operacionais ou entidades de classificação ambígua, garantindo a pureza operacional da amostra.

A Etapa 2 (Filtro de Alvo) removeu 7.551 linhas. Esta "perda" não é um erro, mas sim uma consequência metodológica direta da adoção de um horizonte preditivo de 1 ano (T+4). Conforme será detalhado na Seção 3.2, linhas ao final da série de cada empresa, que não possuíam um alvo futuro conhecido para o treinamento ou validação do modelo, foram corretamente descartadas.

A Etapa 3 (Filtro de volume) removeu 48 empresas que possuíam menos de 16 observações válidas (pós-filtros 1 e 2). O limiar de 16 trimestres (4 anos) não foi arbitrário; ele foi definido empiricamente, alinhando-se ao primeiro quartil (Q1) da distribuição de dados da amostra (que era de 18 observações), garantindo que apenas empresas com histórico insuficiente fossem removidas.

Finalmente, a Etapa 4 (Filtro de viés de seleção) foi crucial. Seguindo as diretrizes metodológicas de Oshiro (2016) e Horta, Alves e Carvalho (2014), foram incluídas apenas empresas que não estavam em situação de dificuldade financeira no início da série histórica, de modo a evitar viés na construção do modelo preditivo. Como o objetivo foi prever eventos de distress e não apenas constatá-los, essa seleção garantiu que o modelo aprendesse padrões de deterioração a partir de condições normais. Esta etapa removeu 143 empresas que não possuíam um baseline saudável de, no mínimo, quatro trimestres (1 ano).

Tabela 2: Distribuição da base por setor

Setor	Empresas	Empresas (%)	Observações	Observações (%)
Consumo cíclico	59	31,1%	2.663	29,5%
Bens industriais	38	20,0%	1.974	21,9%
Utilidade pública	38	20,0%	2.011	22,3%
Materiais básicos	15	7,9%	663	7,3%
Consumo não cíclico	13	6,8%	557	6,2%
Saúde	13	6,8%	553	6,1%
Tecnologia da informação	6	3,2%	216	2,4%
Comunicações	4	2,1%	192	2,1%
Petróleo, gás e biocombustíveis	4	2,1%	196	2,2%

Fonte: Dados da pesquisa.

A Tabela 2 detalha a composição setorial final das 190 empresas e 9.025 observações. A amostra apresenta boa diversificação, com maior concentração nos setores

de Consumo Cíclico, Bens Industriais e Utilidade Pública, que juntos representam 71,1% das empresas.

Tabela 3: Distribuição da base por ano

Ano	Empresas	Observações
2010	130	520
2011	137	548
2012	141	562
2013	148	592
2014	149	590
2015	150	593
2016	151	598
2017	154	613
2018	158	613
2019	158	624
2020	173	684
2021	172	676
2022	168	670
2023	168	661
2024	162	481

Fonte: Dados da pesquisa.

Adicionalmente, a Tabela 3 ilustra a distribuição das observações ao longo do período do estudo (2010-2024). A estabilidade da amostra indica que o dataset não sofre de viés de sobrevivência significativo e captura dados de múltiplos ciclos econômicos.

3.2 Construção da variável dependente *financial distress*

A variável dependente (y) deste estudo é binária (0 ou 1) e indica se uma empresa apresentará um evento de *financial distress* (FD) em um horizonte de previsão de um ano.

A definição do evento de distress foi baseada em critérios contábeis e jurídicos, alinhados com a literatura (Oshiro, 2016; Lunardi et al., 2020), para identificar sinais de deterioração operacional antes da falência formal.

Uma empresa foi classificada como em *financial distress* ($FD = 1$) se atendeu a pelo menos um dos três seguintes critérios, calculados com base nos dados dos últimos doze meses (*Trailing twelve months* - TTM) para mitigar a volatilidade trimestral:

Prejuízo líquido persistente: Ocorrência de lucro líquido negativo (TTM) por dois períodos consecutivos. Este critério identifica uma falha fundamental da empresa em gerar retorno aos acionistas.

Incapacidade de Cobertura de Juros: O indicador de cobertura de juros (TTM) foi inferior a 1 por dois períodos consecutivos, indicando incapacidade persistente de gerar resultado operacional suficiente para cobrir despesas financeiras. Este critério é amplamente utilizado na literatura nacional e internacional para sinalizar desequilíbrio na estrutura de capital, conforme Oshiro (2016) e Lunardi, Bilk e Rodrigues Junior (2020).

Situação de Recuperação Judicial (RJ): Conforme identificação nos Boletins Diários de Informações (BDI) da B3. As empresas foram classificadas como em RJ a partir da primeira menção formal no BDI, sendo posteriormente detalhada a data de ingresso no processo e, se aplicável, a data de saída. Esse procedimento seguiu a prática adotada por Horta, Alves e Carvalho (2014) para caracterização de insolvência com base em eventos jurídicos públicos.

As observações que não atenderam a nenhum desses critérios foram classificadas como $FD = 0$.

Com o objetivo de preservar o rigor preditivo do modelo, foi estabelecido um horizonte de previsão de um ano. A variável dependente (FD) foi defasada para o futuro, garantindo que o modelo utilizasse apenas informações disponíveis no trimestre T para prever a ocorrência de distress em qualquer momento nos quatro trimestres seguintes (T+1 a T+4). Essa abordagem *ex-ante* evita o viés de antecipação e está alinhada a estudos como os de Lunardi et al. (2020), que também utilizaram dados trimestrais para horizontes de um ano.

Para a modelagem final, os três critérios de *distress* foram agregados em uma única variável-alvo binária. Embora a ocorrência de múltiplos critérios simultâneos possam indicar a gravidade do evento, a abordagem binária foi adotada para simplificar a tarefa de classificação e garantir a comparabilidade direta com os modelos tradicionais, que também geram previsões binárias (solvente ou insolvente).

3.3 Engenharia de features e variáveis de controle

Após a definição da amostra final e da variável dependente, a etapa seguinte consistiu na construção do universo de variáveis independentes (preditoras, ou "*features*"). O objetivo foi criar um conjunto de dados robusto, alinhado à literatura, que combinasse indicadores financeiros tradicionais com variáveis de controle e macroeconômicas.

O núcleo do conjunto de *features* foi composto por um amplo espectro de indicadores financeiros, calculados para cada observação empresa-trimestre. Para mitigar a volatilidade de curto prazo e capturar tendências mais estáveis, optou-se por utilizar predominantemente indicadores calculados sobre os dados dos últimos doze meses (TTM, *Trailing Twelve Months*). Esta metodologia consiste na agregação dos valores do trimestre de referência somados aos três trimestres imediatamente anteriores, formando uma janela móvel anual. Essa abordagem é essencial para eliminar os efeitos da

sazonalidade inerentes aos reportes trimestrais, permitindo que cada ponto de dados reflita o desempenho de um ciclo fiscal completo tornando-os comparáveis independentemente do período do ano. Esses indicadores cobriram as dimensões clássicas da análise financeira: Liquidez, endividamento, eficiência, lucratividade e rentabilidade.

Seguindo a proposta deste trabalho, o universo de *features* foi enriquecido com os componentes individuais dos modelos tradicionais revisados na Seção 2.2. Em vez de utilizar apenas os *scores* finais (como ALT ou KAN), variáveis clássicas como A1 de Altman, M2 de Matias e E5 de Elizabetsky foram incluídas como *features* candidatas. Esta abordagem permite que o modelo de machine learning teste o poder preditivo desses "ingredientes" clássicos em uma nova arquitetura, não linear e competitiva.

Adicionalmente, para aprimorar a capacidade preditiva do modelo e reduzir a influência de fatores contextuais, foram incorporadas variáveis de controle, conforme sugerido por Oshiro (2016), Rezende et al. (2017) e Lunardi et al. (2020). As variáveis de controle utilizadas foram:

Setor econômico: Representado por variáveis *dummy* para cada setor principal da B3 permitindo ao modelo capturar riscos setoriais específicos.

Tamanho da empresa: Medido pelo logaritmo natural do Ativo total, conforme Lunardi et al. (2020), para mitigar distorções provocadas por escalas muito diferentes entre empresas.

Contexto Sazonal e Temporal: Variáveis *dummy* para cada trimestre para controlar choques macroeconômicos e mudanças regulatórias ao longo do tempo.

Choque Exógeno: Uma variável *dummy* que assumiu valor 1 entre o primeiro trimestre de 2020 e o último de 2021.

Por fim, foram incorporadas variáveis macroeconômicas trimestrais (calculadas na métrica TTM), com o objetivo de controlar o efeito do ambiente externo sobre a performance financeira. Para garantir que o modelo utilizasse apenas informações publicamente disponíveis no momento da previsão e evitar viés de antecipação (*lookahead bias*), todos os indicadores macroeconômicos foram defasados em um trimestre. Desta forma, para a previsão no trimestre t , foram utilizados os dados macroeconômicos referentes ao trimestre $t-1$. A inclusão dessas variáveis é coerente com estudos como os de Rahayu, Kusuma e Arifin (2025), que demonstraram a relevância de fatores como PIB, inflação, juros e câmbio na previsão de *distress* em mercados emergentes. As variáveis macroeconômicas utilizadas foram taxa de juros (SELIC %)

indicador do custo de capital, inflação (IPCA %) representando o custo de insumos, variação do câmbio (USD/BRL %) medindo a exposição cambial e variação da atividade econômica (PIB %) representando o ciclo econômico.

Este processo resultou em um universo final de 59 *features* candidatas (incluindo indicadores TTM, componentes clássicos e variáveis de controle/macro). Este conjunto de dados abrangente serve como *input* para o *pipeline* de seleção de variáveis e modelagem, detalhado na seção seguinte.

Quadro 3: Variáveis independentes

Tipo	Código	Variável	Fórmula
Dummies	COVID	<i>dummy</i> representando efeito da pandemia de COVID-19	-
	TRI	<i>dummy</i> representando trimestres	-
	SETOR	<i>dummy</i> representando setores	-
	LOG	representação do tamanho da empresa via log de ativos corrigidos a IPCA	-
Macroeconômicas	CAMBIO	representação da variação cambial acumulada	-
	IPCA	representação da inflação acumulada	-
	PIB REAL	representação da produção nacional acumulada	-
	SELIC	representação da variação da taxa de juros	-

Mercado	L/A	lucro por ação	$\frac{\text{Resultado líquido}}{\text{Ações}}$
	P/E	preço sobre lucro	$\frac{\text{Valor de fechamento}}{\text{Resultado por ação}}$
	PBV	preço sobre patrimônio	$\frac{(\text{Valor de fechamento} * \text{Ações})}{\text{Patrimônio líquido}}$
	PM	preço sobre passivo	$\frac{(\text{Valor de fechamento} * \text{Ações})}{\text{Passivo exigível}}$
	PV	preço sobre venda	$\frac{(\text{Valor de fechamento} * \text{Ações})}{\text{Receita}}$
	VE/E	valor empresa por EBIT	$\frac{(\text{Valor de fechamento} * \text{Ações}) + \text{Dívida}}{\text{EBIT}}$
	VE/R	valor empresa por receita	$\frac{(\text{Valor de fechamento} * \text{Ações}) + \text{Dívida}}{\text{Receita}}$
Eficiência	CCX	ciclo de caixa	PMRE + PMRC – PMPF
	COP	ciclo operacional	PMRE + PMRC
	C/REC	custo por receita	$\frac{\text{Custos}}{\text{Receita}}$
	EFREC	eficiência de recebíveis	$\frac{\text{Receita}}{\text{Contas a receber}}$
	GAT	giro do ativo	$\frac{\text{Receita}}{\text{Ativo}}$
	GAT-1	giro do ativo defasado	$\frac{\text{Receita}}{\text{Ativo}(t - 1)}$
	PMPF	prazo médio de pagamento a fornecedores	$\frac{\text{MÉDIA}(\text{Fornecedores})}{\Delta \text{Estoques} + \text{Custos}}$
	PMRC	prazo médio de recebimento de clientes	$\frac{\text{MÉDIA}(\text{Contas a receber})}{\text{Receita}}$
	PRME	prazo médio de renovação de estoques	$\frac{\text{MÉDIA}(\text{Estoques})}{\text{Custos}}$
Endividamento	ALCP	alavancagem de curto prazo	$\frac{\text{Passivo circulante}}{\text{Passivo exigível}}$
	ALLP	alavancagem de longo prazo	$\frac{\text{Passivo não circulante}}{\text{Passivo exigível}}$
	E/A	empréstimos sobre ativo	$\frac{\text{Empréstimos}}{\text{Ativo}}$
	ECP/A	empréstimos de curto prazo sobre ativo	$\frac{\text{Empréstimos de curto prazo}}{\text{Ativo}}$
	ELP/A	empréstimos de longo prazo sobre ativo	$\frac{\text{Empréstimos de longo prazo}}{\text{Ativo}}$
	GE	grau de endividamento	$\frac{\text{Passivo exigível}}{\text{Patrimônio líquido}}$

Liquidez	LGERAL	liquidez geral	$\frac{\text{Ativo circulante} + \text{Realizável LP}}{\text{Passivo exigível}}$
	LIMED	liquidez imediata	$\frac{\text{Ativo circulante}}{\text{Passivo circulante}}$
	LCORR	liquidez corrente	$\frac{\text{Caixa e equivalentes de caixa}}{\text{Passivo circulante}}$
	LSECA	liquidez seca	$\frac{\text{Ativo circulante} - \text{Estoques} - \text{Despesas antecipadas}}{\text{Passivo circulante}}$
Lucratividade	ATBL	ativo de baixa liquidez	$\frac{\text{Estoques} + \text{Despesas antecipadas}}{\text{Passivo circulante}}$
	MBRUT	margem bruta	$\frac{\text{Resultado bruto}}{\text{Receita}}$
	MEBIT	margem EBIT	$\frac{\text{EBIT}}{\text{Receita}}$
	MLIQ	margem líquida	$\frac{\text{Resultado líquido}}{\text{Receita}}$
	RENTJUR	rentabilidade de juros	$\frac{\text{Receitas financeiras}}{\text{Despesas financeiras}}$
Rentabilidade	ROA	<i>return of assets</i>	$\frac{\text{Resultado líquido}}{\text{Ativo}}$
	ROE	<i>return of equity</i>	$\frac{\text{Resultado líquido}}{\text{Patrimônio líquido}}$
	ROI	<i>return of investment</i>	$\frac{\text{Resultado líquido}}{\text{MÉDIA(Investimentos)}}$
	ROIC	<i>return of invested capital</i>	$\frac{\text{Resultado líquido}}{\text{MÉDIA(Capital Investido)}}$
	ROTA	<i>return of total assets</i>	$\frac{\text{Resultado líquido}}{\text{MÉDIA(Ativo)}}$

Modelos tradicionais	A1	A1	$\frac{\text{Ativo circulante} - \text{Passivo circulante}}{\text{Ativo total}}$
	A2	A2	$\frac{\text{Lucros Retidos}}{\text{Ativo total}}$
	A3	A3	$\frac{\text{EBIT}}{\text{Ativo total}}$
	A4~	proxy de A4	$\frac{\text{Patrim\~{o}nio l\~{i}quido}}{\text{Passivo exig\~{i}vel}}$
	E2	E2	$\frac{\text{Caixa e equivalentes}}{\text{Ativo permanente}}$
	E3	E3	$\frac{\text{Contas a receber}}{\text{Ativo total}}$
	E4	E4	$\frac{\text{Estoques}}{\text{Ativo total}}$
	E5	E5	$\frac{\text{Passivo circulante}}{\text{Ativo total}}$
	M1	M1	$\frac{\text{Patrim\~{o}nio l\~{i}quido}}{\text{Ativo total}}$
	M2	M2	$\frac{\text{D\~{i}vida total}}{\text{Ativo circulante}}$
	M3	M3	$\frac{\text{Fornecedores}}{\text{Ativo total}}$
	M5	M5	$\frac{\text{Ativo circulante}}{\text{Passivo circulante}}$
	M6	M6	$\frac{\text{EBIT}}{\text{Resultado bruto}}$
	Z2	Z2	$\frac{\text{Passivo exig\~{i}vel}}{\text{Ativo total}}$

Fonte: Elaborado pelo autor.

Nota: Os modelos tradicionais e suas f\~{o}rmulas foram explorados no quadro 2. Foram entregues ao *pipeline* as que n\~{a}o se repetem em rela\~{c}\~{a}o \~{a}s vari\~{a}veis independentes j\~{a} calculadas.

O Quadro 3 detalha o universo completo das 59 *features* candidatas, agrupadas por categoria anal\~{i}tica e com suas respectivas f\~{o}rmulas conceituais. Conforme mencionado, a maioria dos indicadores financeiros foi calculada com base na m\~{e}trica para estabilizar a s\~{e}rie temporal. Os componentes dos modelos cl\~{a}ssicos (prefixos A, E, M, Z) tamb\~{e}m s\~{a}o listados.

3.4 Pipeline de modelagem e valida\~{c}\~{a}o cruzada

A avalia\~{c}\~{a}o de performance dos modelos foi conduzida por meio de um *pipeline* metodol\~{o}gico integrado, desenhado para garantir rigor, interpretabilidade e prevenir o vazamento de informa\~{c}\~{a}o. A estrat\~{e}gia central, conforme recomenda\~{c}\~{a}o de Horta, Alves e Carvalho (2014), foi a valida\~{c}\~{a}o cruzada temporal (*TimeSeriesSplit* com $k=10$), que

promove uma avaliação robusta da capacidade de generalização dos modelos em dados cronológicos, garantindo que os dados de validação sejam sempre posteriores aos de treino.

Um princípio fundamental norteou este *pipeline*: todas as etapas de pré-processamento que envolvem "aprendizagem" a partir dos dados (como imputação, escalonamento e tratamento de outliers) e a própria seleção de variáveis foram aninhadas. Ou seja, foram executadas iterativamente dentro de cada um dos 10 *folds*, sendo aprendidas exclusivamente sobre os dados de treino daquele *fold* e, em seguida, aplicadas aos dados de validação.

Este *pipeline* aninhado seguiu uma sequência sistemática de sete etapas em cada *fold*:

Primeiramente, foi realizado o pré-processamento dos dados. Esta etapa iniciou com o tratamento de *outliers* extremos via winsorização, onde os valores das *features* contínuas foram limitados aos seus respectivos percentis 1% e 99%. Em seguida, para o tratamento de dados ausentes pontuais, optou-se pela aplicação do algoritmo *KNNImputer*, conforme a abordagem de Horta et al. (2014), que preserva a estrutura multivariada dos dados sem a perda de amostras. Por fim, os dados foram submetidos ao *StandardScaler* para padronização (média 0, desvio 1), um requisito para o *Lasso* e para o *MLP*.

A segunda etapa consistiu na seleção de variáveis explicativas. Seguindo a abordagem híbrida de Zhang et al. (2022), dois métodos *wrapper* foram aplicados nos dados de treino processados: *LassoCV* (para seleção linear) e *Random Forest* com *SelectFromModel* (para seleção não-linear). O conjunto final de *features* para aquele *fold* foi definido pela interseção dos *features* selecionados por ambos os métodos, garantindo um subconjunto de alta relevância e robustez.

A terceira etapa foi o balanceamento de classes. A análise da amostra revelou um desbalanceamento leve (67% Classe 0 vs. 33% Classe 1). Em vez de *SMOTE*, que pode induzir a *overfitting* (Nguyen et al., 2023), optou-se pela técnica de reponderação (*class_weight='balanced'*). Esta abordagem instrui o modelo a aplicar uma penalidade maior aos erros cometidos na classe minoritária (FD=1), resolvendo o desbalanceamento sem criar dados sintéticos.

A quarta etapa foi a modelagem com redes neurais artificiais. Uma arquitetura *feedforward* foi treinada usando as *features* selecionadas e os pesos de classe. O modelo

consistiu em duas camadas ocultas (64 e 32 neurônios, ativação ReLU) e uma camada de saída (1 neurônio, ativação sigmoid), compilado com o otimizador Adam e *loss='binary_crossentropy'*.

O Quadro 4 detalha a arquitetura-base completa e os hiper parâmetros utilizados para o modelo MLP, os quais foram definidos com base em boas práticas da literatura e validados empiricamente nos testes preliminares. A arquitetura foi projetada para equilibrar capacidade de aprendizado (camadas densas com ativação ReLU) e prevenção de overfitting (via *Dropout* e *EarlyStopping*).

Quadro 4: Arquitetura e hiper parâmetros do modelo

Parâmetro	Tipo/Configuração	Justificativa
Camada de Entrada	Input(shape=(N_Features))	N_Features é dinâmico (definido pela seleção em cada fold).
Camada Oculta 1	Dense(64, activation='relu')	Captura de padrões de complexidade primária.
Regularização 1	Dropout(0.2)	Prevenção de <i>overfitting</i> .
Camada Oculta 2	Dense(32, activation='relu')	Refinamento dos padrões.
Regularização 2	Dropout(0.2)	Prevenção de <i>overfitting</i> .
Camada de Saída	Dense(1, activation='sigmoid')	Adequada para classificação binária (probabilidade 0-1).
Otimizador	Adam(learning_rate=0.0001)	Otimizador robusto com taxa de aprendizado controlada.
Função de Perda	binary_crossentropy	Padrão para classificação binária.
Balanceamento	class_weight='balanced'	Compensa o desbalanceamento 67/33.
Parada Antecipada	EarlyStopping(monitor='val_AUC')	Interrompe o treino quando a performance de validação para de melhorar.

Fonte: Elaborado pelo autor.

A quinta etapa foi a prevenção de *overfitting*. A arquitetura do MLP incluiu camadas de *Dropout* (0.2), e o treinamento utilizou *EarlyStopping* (monitorando *val_AUC*) para interromper o processo quando a performance na validação parava de melhorar.

A sexta etapa do *pipeline* consistiu na avaliação comparativa da performance. Para garantir um comparativo academicamente rigoroso, capaz de isolar o verdadeiro impacto preditivo da arquitetura não-linear da rede neural, foi estabelecido um *benchmark* principal baseado em Regressão Logística, sucessora moderna dos modelos estatísticos lineares. Este modelo Logit foi treinado em cada fold utilizando exatamente os mesmos insumos do MLP, ou seja, o mesmo subconjunto de *features* selecionadas e a mesma técnica de reponderação de classes. Adicionalmente, como *baseline* histórico, os *scores* tradicionais foram calculados e avaliados de duas formas distintas: primeiramente,

utilizando seus limiares de corte fixos, conforme a literatura original, e, em segundo, recalibrando um limiar ótimo via curva ROC dentro de cada *fold* de treino, permitindo uma adaptação à amostra atual.

O Quadro 5 resume esta estratégia de avaliação comparativa, contrastando a abordagem de treinamento e avaliação do MLP e do Logit com as duas metodologias aplicadas aos *benchmarks* tradicionais. Esta distinção é crucial: enquanto o MLP e o Logit aprendem seus parâmetros do zero em cada *fold*, os *benchmarks* são testados tanto em sua forma padrão quanto em sua forma otimizada para a amostra atual.

Quadro 5: Estratégia de avaliação comparativa dos modelos

Modelo	Fonte das Features	Estratégia de Treino	Estratégia de Avaliação
MLP	<i>Features</i> selecionadas (Lasso+RF) no <i>fold</i>	Treinado em X_train_resampled (com class_weight)	Testado em X_val_scaled (do <i>fold</i>).
LOGIT	<i>Features</i> selecionadas (Lasso+RF) no <i>fold</i>	Treinado em X_train_resampled (com class_weight)	Testado em X_val_scaled (do <i>fold</i>).
Benchmark (Tradicional)	<i>Scores</i> pré-calculados	Limiar fixo da literatura	Testado em y_val_fold (do <i>fold</i>).
Benchmark (Calibrado)	<i>Scores</i> pré-calculados	Limiar ótimo (via ROC) aprendido <i>apenas</i> em (score train, y train fold).	Testado em score_val (do <i>fold</i>) usando o novo limiar.

Fonte: Elaborado pelo autor.

Finalmente, a sétima etapa foi o registro da performance. A performance de todos os modelos (MLP, Logit e Tradicionais) foi registrada usando um conjunto de métricas de classificação (como F1-Score) e de ranking (como AUC e AUPR). As definições formais, justificativas e fórmulas para cada métrica utilizada são detalhadas na Seção 3.6. Os resultados médios e o desvio padrão dos 10 folds são apresentados e discutidos no Capítulo 4.

Este *pipeline* de validação cruzada (*k-fold*) foi a primeira fase da avaliação, utilizada para medir a estabilidade do modelo (Tabela 4) e identificar o conjunto de *features* robustas (Quadro 6). Para a avaliação de performance final e para a análise de interpretabilidade (SHAP), uma segunda fase de testes foi executada: os modelos (MLP e Logit) foram treinados nos 80% iniciais da amostra (N=7.220) e testados nos 20% de dados mais recentes (N=1.805), que foram temporalmente separados. Esta abordagem de *split* temporal 80/20 simula um cenário de produção real, onde o modelo é treinado no

passado para prever o 'futuro' desconhecido. Os resultados desta avaliação final são apresentados na Seção 4.1 (Tabela 5).

3.5 Interpretabilidade do modelo

Embora as RNAs apresentem alta performance, elas são frequentemente criticadas por sua baixa interpretabilidade (o efeito "caixa-preta"), como destacado por Rudin (2019). Para mitigar essa limitação, foi implementada uma análise de interpretabilidade pós-hoc utilizando *SHAP* (*SHapley Additive exPlanations*).

Para esta análise, um modelo final foi treinado utilizando-se a totalidade da amostra de 9.025 observações. A definição do conjunto final de *features* foi baseada em um critério de estabilidade de seleção (*Selection Stability*), avaliando a consistência com que cada variável foi selecionada durante o *pipeline* de validação cruzada ($k=10$). Foi realizado um cômputo de frequência, e foram retidas apenas as *features* que foram selecionadas em, no mínimo, 6 dos 10 folds. Este limiar foi adotado para garantir que o modelo de interpretabilidade fosse construído apenas com variáveis que demonstraram relevância preditiva consistente através da maioria das partições dos dados, mitigando o risco de sobre ajuste a *features* espúrias.

Conforme recomendado por Tran et al. (2022), foi aplicado o *GradientExplainer* do SHAP, que é otimizado para arquiteturas Keras/TensorFlow. Esta técnica, baseada na teoria dos jogos, calcula a contribuição marginal de cada feature para cada predição individual, permitindo uma análise de interpretabilidade tanto global (quais *features* são mais importantes em média) quanto local (por que uma empresa específica recebeu um *score* de risco). Os resultados desta análise são apresentados no Capítulo 4.

3.6 Métricas de avaliação de performance

Para avaliar e comparar de forma rigorosa a performance dos modelos (MLP, Logit e *benchmarks* tradicionais), este estudo utilizou um conjunto de métricas complementares, divididas em duas categorias: métricas de classificação (dependentes de limiar) e métricas de *ranking* (independentes de limiar).

A base para as métricas de classificação é a Matriz de Confusão, que analisa quatro cenários de previsão. O primeiro é o Verdadeiro Positivo (VP), onde o modelo prevê "*distress*" e a empresa realmente entra em *distress*. O segundo é o Verdadeiro Negativo (VN), onde o modelo prevê "saudável" e a empresa permanece saudável. Os dois cenários de erro são o Falso Positivo (FP), que ocorre quando o modelo prevê "*distress*" mas a

empresa permanece saudável (erro tipo I), e o Falso Negativo (FN), que ocorre quando o modelo prevê "saudável" mas a empresa entra em distress (erro tipo II).

A partir destes quatro resultados, foram calculadas as métricas de classificação. A Precisão, calculada como $\frac{VP}{VP+FP}$, mede a acurácia das previsões positivas, respondendo quantas empresas classificadas como em risco realmente entraram em distress. O *recall*, calculado como $\frac{VP}{VP+FN}$, mede a capacidade do modelo em identificar todos os casos positivos, ou seja, quantas das empresas que realmente entraram em distress o modelo conseguiu prever, sendo o Falso Negativo o erro de maior custo na gestão de risco. Por fim, o *F1-Score*, sendo a média harmônica entre precisão e *recall* $\left(2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}}\right)$, foi usado como a métrica principal para encontrar o equilíbrio ótimo entre os dois, sendo fundamental para a calibração do limiar de decisão final dos modelos.

Para avaliar a performance geral dos modelos independentemente de um limiar de corte específico, foram usadas métricas de ranking. A AUC (*Area Under the ROC Curve*) é a métrica padrão para medir o poder de discriminação de um modelo, avaliando sua capacidade em atribuir uma pontuação de risco maior a uma empresa que entrará em distress do que a uma empresa saudável (onde 0.5 é aleatório e 1.0 é perfeito). Complementarmente, a AUPR (*Area Under the Precision-Recall Curve*) foi incluída. O AUPR é considerado mais robusto e fidedigno que o AUC em datasets com classes desbalanceadas (RAHMAN; ZHU, 2024). Dado que a classe "*distress*" é minoritária, o AUPR garante que a alta performance do AUC não seja uma ilusão estatística causada pela maioria de casos "saudáveis".

4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os resultados empíricos obtidos a partir da execução da metodologia descrita no Capítulo 3. A análise está dividida em duas seções centrais. A Seção 4.1 detalha a performance comparativa dos modelos, respondendo à pergunta de pesquisa ao contrastar a Rede Neural (MLP) com o *benchmark* linear (Regressão Logística) e os modelos tradicionais. A Seção 4.2 aprofunda a análise, utilizando a técnica SHAP para comparar a interpretabilidade dos modelos MLP e Logit, explicando quais fatores foram determinantes para seus respectivos resultados.

4.1 Performance comparativa dos modelos

A avaliação de performance foi conduzida em um *pipeline* de duas fases para garantir tanto a estabilidade quanto a precisão preditiva em dados recentes. A primeira

fase consistiu em uma validação cruzada temporal ($k=10$, via *TimeSeriesSplit*) para avaliar a robustez geral do *pipeline* e dos modelos. A Tabela 4 consolida a média e o desvio padrão dos resultados obtidos nos 10 *folds*.

Tabela 4: Métricas da validação cruzada K-Fold

modelo	AUC		AUPR		F1		Recall		Precisão		
	média	desvio p	média	desvio p	média	desvio p	média	desvio p	média	desvio p	
	MLP	0,8728	0,0708	0,8104	0,1607	0,6811	0,1849	0,6546	0,2452	0,7764	0,1102
	LOGIT	0,8756	0,0399	0,8115	0,1236	0,6849	0,1604	0,6627	0,2191	0,7631	0,1193
ORIGINAL	ALT	0,6180	0,0886	0,4863	0,1498	0,5170	0,1063	0,5620	0,1010	0,4876	0,1302
	ALT~	0,7478	0,0633	0,6340	0,1449	0,6006	0,0819	0,7378	0,0918	0,5227	0,1241
	ELI	0,7526	0,0506	0,6507	0,1492	0,6036	0,1015	0,5946	0,1056	0,6225	0,1269
	KAN	0,6271	0,0477	0,4819	0,1339	0,0771	0,0600	0,0416	0,0329	0,6265	0,3468
	MAT	0,6475	0,0421	0,5311	0,1280	0,4789	0,0882	0,4887	0,0672	0,4850	0,1378
	ZMI	0,5785	0,0573	0,4686	0,1418	0,3056	0,1295	0,2379	0,1165	0,4840	0,2126
AJUSTADO	ALT	0,6180	0,0886	0,4863	0,1498	0,5305	0,1170	0,9827	0,0433	0,3717	0,1106
	ALT~	0,7478	0,0633	0,6340	0,1449	0,5765	0,1049	0,8632	0,1044	0,4462	0,1161
	ELI	0,7526	0,0506	0,6507	0,1492	0,5984	0,1127	0,5806	0,1250	0,6261	0,1262
	KAN	0,6271	0,0477	0,4819	0,1339	0,5275	0,1034	0,8767	0,0567	0,3912	0,1221
	MAT	0,6475	0,0421	0,5311	0,1280	0,4988	0,1943	0,7488	0,3041	0,4802	0,2054
	ZMI	0,5785	0,0573	0,4686	0,1418	0,4842	0,1966	0,8623	0,3230	0,3418	0,1513

Fonte: Dados da pesquisa.

Nota: **ORIGINAL** = Modelos tradicionais originais sem balanceamento de limiares; **AJUSTADO** = Modelos tradicionais ajustados com balanceamento de limiar para base de dados; **MLP** = Arquitetura da Rede Neural - *Perceptron* de múltiplas camadas; **LOGIT** = Modelo de regressão logística utilizado como benchmark; **ALT** = Modelo histórico de Altman; **ALT~** = Proxy do modelo de Altman usando patrimônio líquido no lugar de valor de mercado, simulando análise em empresas de capital fechado; **ELI** = Modelo histórico de Elizabetsky; **KAN** = Modelo histórico de Kanitz; **MAT** = Modelo histórico de Matias; **ZMI** = Modelo histórico de Zmijewski.

A análise da Tabela 4 revela a primeira descoberta fundamental: a engenharia de *features* é o principal motor de performance. Os modelos tradicionais que serviram de *baseline*, atingiram um teto de performance de 0.7526 AUC. Em contraste, os dois modelos modernos (MLP e Logit), treinados com o conjunto de *features* híbrido, atingiram um patamar de performance vastamente superior e estatisticamente idêntico, com o Logit (AUC 0.8756) e o MLP (AUC 0.8728) apresentando um empate técnico. O baixo desvio padrão (0.0399-0.0708) em ambos também confirma a estabilidade do *pipeline*.

Para identificar quais fatores foram responsáveis por este salto de performance, foi analisada a consistência da seleção de *features* dentro dos 10 *folds*. O Quadro 6 detalha a frequência com que cada feature foi selecionada.

Quadro 6: Consistência das features

Código	Contagem de seleções	Conclusão
ECP/A, A3	10 de 10	Robusto
IPCA, ROTA	9 de 10	Secundário
M5	7 de 10	Secundário
GAT-1, ROE, MBRUT, LOG, GE, PIB	6 de 10	Secundário
ROIC, PMRC, E/A	5 de 10	Instável
ROI, COP, MLIQ, E3	4 de 10	Instável
C/REC, M6, A2, RENTJUR	3 de 10	Instável
SELIC, ROA, PMPF, L/A, M1, Z2	2 de 10	Instável
E2, MEBIT, PRME, CAMBIO, EFREC, VE/E, E5	1 de 10	Instável

Fonte: Dados da pesquisa.

Nota: **ECP/A** = Endividamento de curto prazo sobre ativo; **A3** = Variável 3 do modelo de Altman; **M5** = Variável 5 do modelo de Matias; **GAT-1** = Giro do ativo, com ativo defasado em 1 período; **MBRUT** = Margem bruta; **LOG** = Logaritmo do ativo total, usado para controle de tamanho da empresa; **GE** = Grau de endividamento; **E/A** = Endividamento total sobre ativo; **COP** = Ciclo Operacional; **MLIQ** = Margem líquida; **E3** = Variável 3 do modelo de Elizabetsky; **C/REC** = Custo por receita; **M6** = Variável 6 do modelo de Matias; **A2** = Variável 2 do modelo de Altman; **RENTJUR** = Rentabilidade de juros, **L/A** = Lucro líquido por ação; **M1** = Variável 1 do modelo de Matias; **Z2** = Variável 2 do modelo de Zmijewski; **E2** = Variável 2 do modelo de Elizabetsky; **MEBIT** = Margem EBIT; **EFREC** = Eficiência de recebíveis; **VE/E** = Valor empresa por EBIT; **E5** = Variável 5 do modelo de Elizabetsky.

O Quadro 6 fornece a prova empírica da tese deste trabalho. O conjunto de *features* mais robusto (selecionado em 60% ou mais dos *folds*) é um "conjunto híbrido", validando a discussão da revisão de literatura. A performance foi impulsionada por uma combinação de: (1) Rentabilidade Clássica (A3, ROTA, ROE, M5 e MBRUT), (2) Endividamento (ECP/A, GE), (3) Eficiência (GAT-1), (4) Contexto Macroeconômico (IPCA, PIB) e (5) Controle (LOG). O fato de ambos os modelos (MLP e Logit) terem usado este mesmo pool de variáveis explica o empate técnico na Tabela 4.

Concluída a análise de estabilidade e *features*, a segunda fase da avaliação consistiu no teste final e mais rigoroso: os modelos (MLP e Logit) foram treinados nos 80% iniciais dos dados e testados no conjunto de validação temporalmente separado (os 20% de dados mais recentes, N=1.805). A Tabela apresenta os resultados deste teste.

Tabela 5: Métricas da validação final entre os dois modelos

modelo	AUC	AUPR	F1	Recall	Precisão
MLP	0,9229	0,8642	0,7491	0,7033	0,8012
LOGIT	0,9070	0,8453	0,7418	0,7068	0,7803

Fonte: Dados da pesquisa

Nota: **MLP** = Arquitetura da Rede Neural - *Perceptron* de múltiplas camadas; **LOGIT** = Modelo de regressão logística utilizado como benchmark.

Os resultados da Tabela 5 revelam a descoberta final: a arquitetura do modelo importa para a otimização. Na comparação direta de arquiteturas (Linear vs. Não-Linear) nos dados mais recentes, o modelo MLP demonstrou uma superioridade estatisticamente significativa. O MLP alcançou um **AUC de 0.9229**, superando o Logit (AUC **0.9070**). Esta vantagem de 1.6 p.p. em AUC sugere que, embora o modelo linear tenha capturado a maior parte do sinal preditivo, a capacidade do MLP de modelar relações não lineares e interações complexas entre elas foi o que garantiu a performance superior nos dados mais recentes.

A análise das métricas de classificação reforça esta conclusão. O MLP atingiu um F1-Score de **0.7491** com uma precisão de **0.8012**, superando o Logit (F1 **0.7418**, precisão **0.7803**). Ambos os modelos provaram ser robustos e equilibrados, identificando corretamente cerca de 70% dos casos reais de *distress* (*recall*) no conjunto de validação.

Adicionalmente, uma análise crítica sob a ótica da gestão de risco atenção especial ao erro tipo II (classificar uma empresa insolvente como saudável), que representa o custo mais severo para credores. Neste estudo, essa dimensão é capturada pela métrica de *recall*. O modelo MLP deste trabalho conseguiu aliar a mitigação desse erro crítico a uma discriminação global superior, demonstrando uma evolução na capacidade de generalização da arquitetura.

Por fim, foi analisada a generalização da performance dos modelos finais através dos diferentes setores presentes no conjunto de validação. A Tabela 6 detalha esta performance setorial.

Tabela 6: Performance por setor

Setor	Observações	Classes presentes	AUC Logit	AUC MLP
Bens industriais	584	2	0,9069	0,9107
Consumo cíclico	421	2	0,9168	0,9167
Utilidade pública	214	2	0,9720	0,9690
Materiais básicos	168	2	0,9251	0,9197
Consumo não cíclico	118	2	0,8388	0,8331
Comunicações	105	2	0,9382	0,9247
Tecnologia da informação	105	2	0,8993	0,9111
Petróleo gás e biocombustíveis	90	1	-	-

Fonte: Dados da pesquisa;

A análise setorial reforça a robustez do *pipeline* híbrido. Em todos os setores com dados suficientes para o cálculo, ambos os modelos mantiveram um poder preditivo muito elevado, com AUCs consistentemente acima de 0.83. Os resultados também demonstram que a vantagem marginal do MLP na amostra agregada não é homogênea. Em setores como Bens industriais e Tecnologia da informação, o MLP foi superior. Em outros, como Utilidade pública e Consumo cíclico, o Logit manteve uma performance igual ou superior, sugerindo que o ganho de não-linearidade do MLP pode ser mais relevante em alguns contextos operacionais do que em outros.

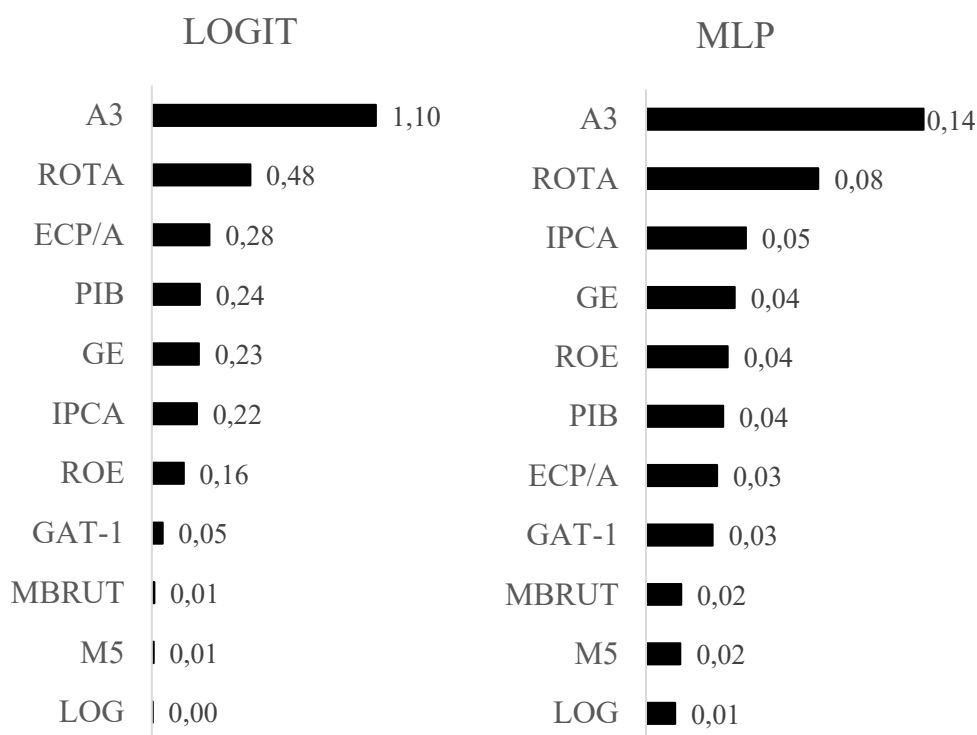
Este padrão de superioridade da rede neural é consistente com a literatura internacional recente. Elhoseny et al. (2022), por exemplo, também encontraram uma vantagem clara das Redes Neurais (Acurácia 0.9134) sobre a Regressão Logística (Acurácia 0.7971). Irvan (2024) e Nguyen et al. (2023) corroboram esta tese, concluindo qualitativamente que as arquiteturas de *deep learning* são "consistentemente os modelos mais eficazes" por acessarem padrões de forma mais eficaz que modelos clássicos. Curiosamente, Sethi e Mahadik (2025) encontraram um empate técnico (Logit \approx RNA) em dados da Índia, similar ao encontrado na validação cruzada (Tabela 4) deste estudo, o que reforça que a vantagem do MLP pode depender do contexto e da complexidade das interações nos dados.

A contribuição mais significativa, no entanto, é contextualizada pelo trabalho de Barboza e Altman (2024). Em seu estudo na América Latina, eles alcançaram um AUC de aproximadamente 0.90 com *Random Forest* (não-linear), mas sem incluir variáveis macroeconômicas, que listaram como uma limitação para "trabalhos futuros". Este trabalho preenche essa lacuna: ao incorporar *features* macro, o modelo MLP deste estudo atingiu um AUC superior (0.9229), demonstrando o valor agregado desses indicadores.

Conclui-se, portanto, que a pergunta de pesquisa central do trabalho foi respondida: a Rede Neural Artificial, quando fundamentada em uma engenharia de *features* robusta que inclui fatores macroeconômicos, é a ferramenta de melhor performance para a previsão de *financial distress* na B3. A seção seguinte detalha como cada modelo utilizou essas *features* para chegar a suas conclusões.

4.2 Análise de interpretabilidade do modelo SHAP

A Seção 4.1 demonstrou que, embora a engenharia de *features* tenha elevado ambos os modelos modernos a um patamar de alta performance ($AUC > 0.90$), a arquitetura do MLP garantiu uma vantagem final sobre a Regressão Logística ($AUC\ 0.9229\ vs\ 0.9070$). Esta seção responde à pergunta metodológica crucial: por quê? Para mitigar a crítica da "caixa-preta" (Rudin, 2019) e comparar os fatores de decisão de ambos os modelos, foi realizada uma análise de interpretabilidade pós-hoc utilizando SHAP (*SHapley Additive exPlanations*). A Figura 1 apresenta a importância global de cada feature, calculada como o valor médio absoluto das contribuições SHAP, e seus respectivos *rankings* de importância para cada modelo.



Fonte: Dados da pesquisa.

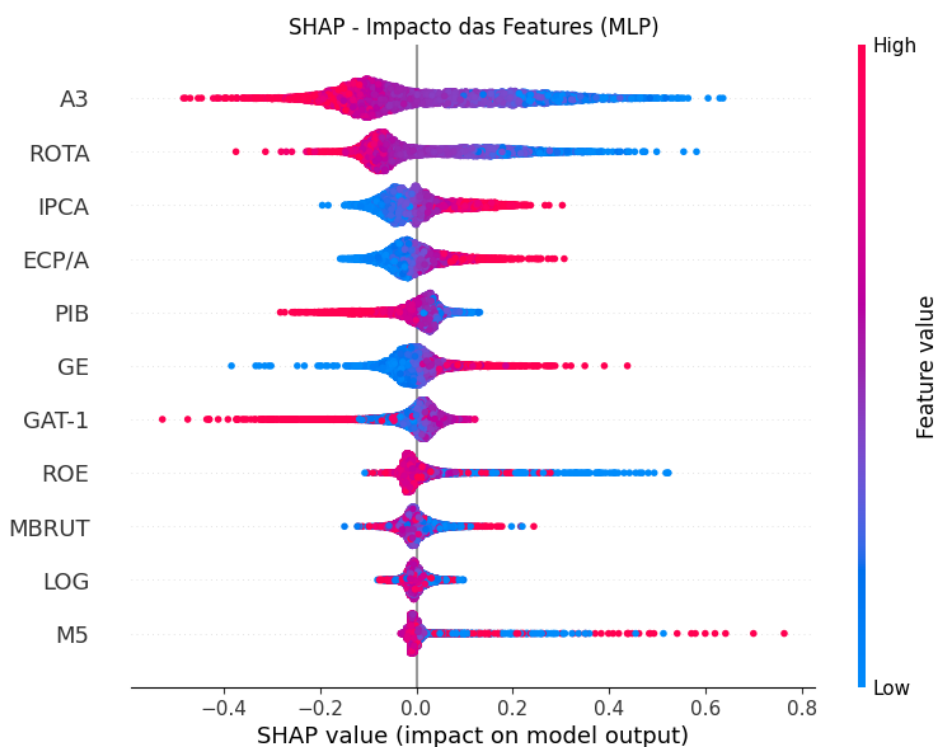
Nota: **ECP/A** = Endividamento de curto prazo sobre ativo; **A3** = Variável 3 do modelo de Altman; **M5** = Variável 5 do modelo de Matias; **GAT-1** = Giro do ativo, com ativo defasado em 1 período; **MBRUT** = Margem bruta; **LOG** = Logaritmo do ativo total, usado para controle de tamanho da empresa; **GE** = Grau de endividamento.

Figura 1: Comparativo entre importância das variáveis para cada modelo

A análise da Figura 1 revela a descoberta mais importante da interpretabilidade: os modelos concordam amplamente sobre o que é mais importante. As duas *features* de maior impacto em ambos os modelos são A3 e ROTA. Esta alta correlação nos *rankings* explica por que o modelo Logit teve uma performance tão forte (AUC 0.9070): os principais impulsionadores do *financial distress* nesta amostra são, de fato, fortemente lineares e ligados à rentabilidade operacional.

Este conjunto de *features* mais importantes também valida empiricamente a tese central do trabalho: a eficácia do modelo híbrido. A presença de duas variáveis macroeconômicas, com alta importância tanto para o Logit quanto para o MLP, confirma a premissa de que negligenciar o ambiente macroeconômico resulta em uma perda significativa de poder preditivo.

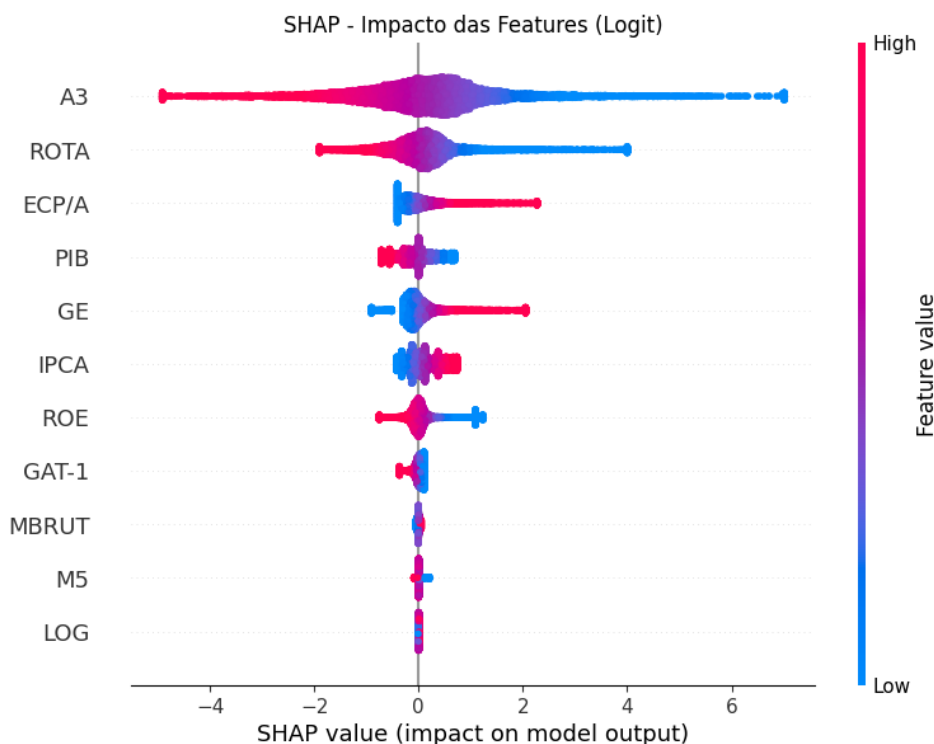
Finalmente, a análise SHAP explica a vantagem de 1.6 p.p. do MLP. Enquanto o Logit só pode avaliar o impacto de A3 de forma isolada, o MLP (Figura 2) captura interações complexas. O gráfico de *beeswarm* do MLP demonstra como o valor de uma *feature* pode ter seu impacto alterado por outras *features*. A Figura 3 em contraste, mostra o impacto mais linear e "limpo" do Logit. Conclui-se que a engenharia de *features* foi a fundação do sucesso, mas a arquitetura da Rede Neural foi o refinamento final, extraindo valor de interações não-lineares que o modelo linear não pôde capturar.



Fonte: Dados da pesquisa.

Nota: **ECP/A** = Endividamento de curto prazo sobre ativo; **A3** = Variável 3 do modelo de Altman; **M5** = Variável 5 do modelo de Matias; **GAT-1** = Giro do ativo, com ativo defasado em 1 período; **MBRUT** = Margem bruta; **LOG** = Logaritmo do ativo total, usado para controle de tamanho da empresa; **GE** = Grau de endividamento.

Figura 2: Importância das features para o modelo MLP



Fonte: Dados da pesquisa.

Nota: **ECP/A** = Endividamento de curto prazo sobre ativo; **A3** = Variável 3 do modelo de Altman; **M5** = Variável 5 do modelo de Matias; **GAT-1** = Giro do ativo, com ativo defasado em 1 período; **MBRUT** = Margem bruta; **LOG** = Logaritmo do ativo total, usado para controle de tamanho da empresa; **GE** = Grau de endividamento.

Figura 3: Importância das features para o modelo LOGIT

5 CONSIDERAÇÕES FINAIS

Este trabalho de conclusão de curso teve como objetivo central avaliar a aplicação e a eficácia de Redes Neurais Artificiais (MLP) na previsão de *financial distress* em empresas listadas na B3, com um horizonte preditivo de um ano. A metodologia buscou responder se arquiteturas não lineares ofereceriam performance superior aos modelos estatísticos tradicionais (tradicionais e lineares), especialmente quando alimentadas por um conjunto híbrido de *features* financeiras, macroeconômicas e de controle.

As conclusões obtidas confirmam à pergunta de pesquisa, mas com uma nuance fundamental que representa a principal contribuição deste estudo. A performance preditiva não é primariamente impulsionada pela complexidade da arquitetura, mas sim pela qualidade da engenharia de *features*. A análise da Seção 4.1 demonstrou que os

modelos tradicionais são insuficientes (AUC 0.5785-0.7526), mas que um modelo linear simples (Regressão Logística), quando treinado com o conjunto de *features* híbrido, atinge um patamar de alta performance (AUC 0.9070). Isso prova que a inclusão de variáveis contextuais, como o IPCA e o PIB, é o fator determinante para o sucesso da previsão.

No entanto, o estudo também validou que, uma vez estabelecida essa base robusta de *features*, a arquitetura não-linear (MLP) oferece uma vantagem de performance significativa, embora marginal. O MLP alcançou um AUC de 0.9229 na validação final, superando o Logit. A análise de interpretabilidade explicou o porquê: ambos os modelos concordaram sobre quais eram as *features* mais importantes (como A3 e ROTA), mas o MLP foi capaz de extrair valor adicional de interações não-lineares entre elas, que o modelo linear, por definição, não pôde capturar.

Sob a ótica econômico-financeira, os achados permitem concluir que o *financial distress* na B3 não é um evento puramente idiossincrático. A predominância de indicadores de rentabilidade operacional combinados a fatores macroeconômicos no topo do ranking de importância sugere que a insolvência no Brasil é frequentemente o resultado da colisão entre ineficiência operacional interna e volatilidade econômica externa. Empresas com baixa geração de caixa operacional tornam-se incapazes de absorver choques de custo de capital ou inflação, validando a necessidade de que modelos de risco incorporem o cenário macroeconômico como fator estrutural, e não apenas acessório.

Do ponto de vista prático, este trabalho oferece um *pipeline* metodológico robusto e um conjunto de 11 *features* validadas que podem ser utilizadas por analistas, gestores e investidores para a construção de modelos de *rating* e *early warning* no mercado brasileiro. A descoberta de que um Logit bem alimentado é "bom o suficiente" oferece um caminho de menor complexidade, enquanto o MLP representa o estado da arte em precisão.

Apesar dos resultados robustos, este estudo possui limitações. A definição de *financial distress* foi baseada em três critérios (Prejuízo, Cobertura de Juros e RJ), mas outras definições (ex: baseadas em valor de mercado ou liquidez de ações) poderiam gerar resultados distintos. Além disso, o *pipeline* de *features* não incluiu dados não estruturados (como análise de sentimento em relatórios da administração) ou indicadores de governança corporativa, que poderiam agregar valor preditivo. Por fim, a arquitetura do

MLP testada foi um *Perceptron* de Múltiplas Camadas (MLP) *feedforward*; modelos sequenciais (como LSTMs ou GRUs) não foram explorados.

Para trabalhos futuros, sugerem-se quatro caminhos principais: (1) a integração de dados não estruturados ao conjunto de *features*; (2) a aplicação de modelos sequenciais (LSTMs) para avaliar se a ordem temporal das *features* importa mais do que o snapshot no trimestre T; (3) a aplicação desta metodologia em um *dataset* de empresas de capital fechado; e (4) a realização de estudos cross-country para avaliar o grau de influência das variáveis macroeconômicas na previsão de distress em diferentes economias, comparando mercados emergentes e desenvolvidos.

REFERÊNCIAS

BARBOSA, A. A. S.; NOBRE, F. C. Mapeamento da produção científica sobre os modelos preditivos de insolvência no Brasil. **Revista Gestão e Secretariado (GeSec)**, São Paulo, v. 14, n. 1, p. 645-661, 2023. Disponível em: <http://dx.doi.org/10.7769/gesec.v14i1.1540>. Acesso em: 9 nov. 2025

BARBOZA, F.; ALTMAN, E. Predicting financial distress in Latin American companies: A comparative analysis of logistic regression and random forest models. **The North American Journal of Economics and Finance**, [S. l.], v. 72, 102158, 2024. Disponível em: <https://doi.org/10.1016/j.najef.2024.102158>. Acesso em: 9 nov. 2025.

BELLOVARY, J. L.; GIACOMINO, D.; AKERS, M. D. A Review of Bankruptcy Prediction Studies: 1930-Present. **Journal of Financial Education**, [S. l.], v. 33, p. 1-42, Winter 2007. Disponível em: https://epublications.marquette.edu/account_fac/25/. Acesso em: 9 nov. 2025.

ELHOSENY, M.; METAWA, N.; SZTANO, G.; EL-HASNONY, I. M. Deep Learning-Based Model for Financial distress Prediction. **Annals of Operations Research**, [S.l.], v. 345, p. 885-907, 2025. Disponível em: <https://doi.org/10.1007/s10479-022-04766-5>. Acesso em: 9 nov. 2025.

HORTA, R. A. M.; ALVES, F. J. S.; CARVALHO, F. A. A. SELEÇÃO DE ATRIBUTOS NA PREVISÃO DE INSOLVÊNCIA: APLICAÇÃO E AVALIAÇÃO USANDO DADOS BRASILEIROS RECENTES. **RAM, Rev. Adm. Mackenzie**, São Paulo, v. 15, n. 1, p. 125-151, jan./fev. 2014. Disponível em: <https://doi.org/10.1590/S1678-69712014000100006>. Acesso em: 9 nov. 2025.

IRVAN, F. M. Comparative analysis of machine learning and deep learning models integrated with Altman Z-score for financial distress prediction in companies listed on the Indonesia Stock Exchange (IDX). **EKOMBIS REVIEW: Jurnal Ilmiah Ekonomi Dan Bisnis**, [S. l.], v. 12, n. 2, p. 2259-2278, abr. 2024. Disponível em: <https://doi.org/10.37676/ekombis.v12i2>. Acesso em: 9 nov. 2025.

LUNARDI, M. A.; BILK, Â.; RODRIGUES JUNIOR, M. M. FINANCIAL DISTRESS E O DESEMPENHO ECONÔMICO, DE MERCADO E FINANCEIRO. **Contabilidade Vista & Revista**, [S. l.], v. 31, n. 1, p. 85-104, 2020. Disponível em: <https://revistas.face.ufmg.br/index.php/contabilidadevistaerevista/article/view/4853>. Acesso em: 9 nov. 2025.

NGUYEN, M.; NGUYEN, B.; LIEU, M. L. Corporate financial distress prediction in a transition economy. **Journal of Forecasting**, [S. l.], v. 43, n. 8, p. 3128-3160, nov. 2024. Disponível em: <https://doi.org/10.1002/for.3177>. Acesso em: 9 nov. 2025.

OSHIRO, R. K. Estruturas de governança corporativa e financial distress: há relação entre conselho de administração e empresas em financial distress? 2016. 107 f. Dissertação (Mestrado Profissional em Finanças e Economia) – Escola de Economia de São Paulo, Fundação Getulio Vargas, São Paulo, 2016.

RAHAYU, N. E. E.; KUSUMA, H.; ARIFIN, Z. The influence of macroeconomic factors on financial distress of companies in Asean: an analysis of SDG implications. **Journal of Lifestyle and SDGs Review**, [S. l.], v. 5, n. 3, p. e05060, 2025. Disponível em: <https://doi.org/10.47172/2965-730X.SDGsReview.v5.n03.pe05060>. Acesso em: 9 nov. 2025.

RAHMAN, M. J.; ZHU, H. Predicting financial distress using machine learning approaches: Evidence China. **Journal of Contemporary Accounting & Economics**, [S. l.], v. 20, n. 1, 100403, 2024. Disponível em: <https://doi.org/10.1016/j.jcae.2024.100403>. Acesso em: 9 nov. 2025.

REZENDE, F. F.; MONTEZANO, R. M. S.; OLIVEIRA, F. N.; LAMEIRA, V. J. Predicting financial distress in publicly-traded companies. **Revista Contabilidade & Finanças**, São Paulo, v. 28, n. 75, p. 390-406, set./dez. 2017. Disponível em: <https://doi.org/10.1590/1808-057x201704460>. Acesso em: 9 nov. 2025.

RUDIN, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. **Nature Machine Intelligence**, [S. l.], v. 1, p. 206-215, maio 2019. Disponível em: <https://doi.org/10.1038/s42256-019-0048-x>. Acesso em: 9 nov. 2025.

SETHI, S. R.; MAHADIK, D. A. Forecasting financial distress for organizational sustainability: An empirical analysis. **Sustainable Futures**, [S. l.], v. 9, 100429, 2025. Disponível em: <https://doi.org/10.1016/j.sftr.2024.100429>. Acesso em: 9 nov. 2025.

SOARES, C. S.; MARIN, L. O.; SANTOS, E. A. CARACTERÍSTICAS DAS PESQUISAS COM APLICAÇÃO DE MODELOS DE PREVISÃO DE INSOLVÊNCIA: UM ESTUDO BIBLIOMÉTRICO NO CONTEXTO BRASILEIRO. **Revista Eletrônica do Alto Vale do Itajaí – REAVI**, v. 10, n. 16, p. 61-80, ago. 2021. Disponível em: <https://doi.org/10.5965/2316419010152021061>. Acesso em: 9 nov. 2025.

TRAN, K. L.; LE, H. A.; NGUYEN, T. H.; NGUYEN, D. T. Explainable machine learning for financial distress prediction: Evidence from Vietnam. **Data**, [S. l.], v. 7, n. 11, 160, 2022. Disponível em: <https://doi.org/10.3390/data7110160>. Acesso em: 9 nov. 2025.

ZHANG, Z.; WU, C.; QU, S.; CHEN, X. An explainable artificial intelligence approach for financial distress prediction. **Information Processing & Management**, [S. l.], v. 59, n. 4, 102988, 2022. Disponível em: <https://doi.org/10.1016/j.ipm.2022.102988>. Acesso em: 9 nov. 2025.