



Universidade do Minho

Escola de Economia e Gestão

Patrícia da Santa Martins Sanches

**Determinantes da Insolvência
das Empresas Cotadas da Zona Euro**



Universidade do Minho

Escola de Economia e Gestão

Patrícia da Santa Martins Sanches

**Determinantes da Insolvência
das Empresas Cotadas da Zona Euro**

Dissertação de Mestrado

Mestrado em Gestão - Especialização em Gestão Geral

Trabalho realizado sob a orientação da
Professora Doutora Sónia Silva

maio 2023

Direitos de Autor e Condições de Utilização do Trabalho por Terceiros

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositórioUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho



Atribuição-NãoComercial-SemDerivações

CC BY-NC-ND

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Agradecimentos

Agradeço a quem tornou possível que esta importante etapa da minha vida fosse concluída, por todo o apoio, sem nunca duvidarem de mim.

Começo por deixar um agradecimento muito especial, à minha orientadora Professora Doutora Sónia Maria da Silva Faria Nogueira da Silva, pelo seu conhecimento, disponibilidade e preocupação que foi imprescindível para a realização desta dissertação.

Aos meus pais, que são as pessoas mais importantes da minha vida. Por sempre me apoiarem, por todo o esforço e por tudo aquilo me passaram ao longo destes anos, devo-vos o que sou e serei eternamente grata. Sem vocês, não seria possível.

Ao meu irmão, que é o meu companheiro de todas as horas e que está sempre ao meu lado. Que seja sempre assim.

Aos meus avós, por tudo o que fazem por mim e por todos os sábios conselhos que partilharam comigo ao longo da vida que, sem dúvida, recordarei para sempre. À minha família, que foi e será sempre o meu suporte, um enorme obrigada por tudo.

Aos meus amigos de infância e aos amigos que tive a sorte de encontrar ao longo do meu caminho, sou grata por serem incansáveis, por sempre me apoiarem e por todos os momentos que partilhamos. Um obrigada não chega, guardo-vos para sempre no coração.

Levo, para a vida, um bocadinho de todos.

Declaração de Integridade

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

Resumo

Os modelos de previsão de insolvência são um instrumento valioso a nível económico e financeiro, uma vez que, para além de permitirem que as empresas avaliem a sua situação financeira, funcionam como um meio de suporte e apoio à forma como as mesmas adotam a sua estratégia e gerem a sua atividade. Neste sentido, revelam a sua importância por representarem uma ferramenta que auxilia os agentes económicos na tomada de decisão no que respeita a anteciparem possíveis cenários de insolvência prevenindo, dentro do possível, a ocorrência de tais eventos.

Neste âmbito, esta dissertação propõe um modelo logístico capaz de prever a insolvência das empresas cotadas da Zona Euro, que revela capacidade preditiva até cinco anos antes do acontecimento. Os determinantes da situação de insolvência incluídos no modelo final são rácios económico-financeiros, testados com recurso à Análise Discriminante Múltipla, usada como análise preliminar, e à Análise Logística, principal método de análise.

A amostra deste estudo é constituída por empresas que se encontram em estado de insolvência e por empresas saudáveis, localizadas na Zona Euro e cotadas em Bolsa de Valores Mobiliários, compreendendo o horizonte temporal entre 2006 até 2021, inclusive. Os dados foram recolhidos da *ORBIS Europe* da *Moody's Analytics* e da *Refinitiv DataStream* e *Worldscope*.

O modelo logístico final de previsão de insolvência é composto pelas variáveis que revelaram maior capacidade preditiva no que diz respeito à deteção da insolvência, nomeadamente: Rentabilidade das Vendas, Valor de Mercado das Vendas, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão. Posteriormente, este modelo foi sujeito a testes de robustez, revelando uma boa capacidade preditiva, com um nível de precisão associado de 80% até cinco anos antes de ser declarada a situação de insolvência.

Palavras-chave: Análise Discriminante Múltipla; Análise Logística; Insolvência; Modelos de Previsão de Insolvência.

Abstract

Financial distress prediction models are a valuable tool at the economic and financial level since, in addition to allowing companies to assess their financial situation, they function as a means of support of corporations' strategy and operating activities management. In this sense, these models reveal their importance by representing a tool that helps economic agents in decision-making with regard to anticipate possible insolvency scenarios, preventing, as far as possible, the occurrence of such events.

In this context, this dissertation proposes a logistic model capable of predicting financial distress of listed companies in the Eurozone, providing predictive ability up to five years before the event occurs. The insolvency situation determinants included in the final distress model are economic and financial ratios, tested using Multiple Discriminant Analysis as a preliminary analysis, and Logistic Analysis as the main method of analysis.

The sample used in this study is formed of insolvent companies and healthy companies from the Eurozone and listed on stock exchange markets, covering the 2006 to 2021 period, *inclusive*. The data was collected from *Moody's Analytics' ORBIS Europe* and *Refinitiv DataStream* and *Worldscope*.

The final logistic model for forecasting financial distress is composed of variables that have revealed the greatest predictive ability regarding insolvency prediction, namely: Return on Sales, Market Value of Sales, Leverage, Solvency, Retained Earnings, and Size. Furthermore, this model passed through different robustness tests, revealing a good predictive capacity with an associated precision level of 80%, up to five years before the insolvency situation.

Keywords: Financial Distress Prediction Models; Insolvency; Logistics Analysis; Multiple Discriminant Analysis.

Índice

1. Introdução	1
2. Revisão da Literatura	5
3. Metodologia	17
4. Descrição da Amostra e Variáveis	20
4.1. Amostra	20
4.2. Variáveis	21
4.3. Estatística Descritiva	24
4.4. Análise de Correlação de <i>Pearson</i>	27
4.5. Análise Univariada	29
5. Análise Multivariada	31
5.1. Análise Discriminante Múltipla	31
5.2. Análise Logística	36
6. Apresentação e Discussão dos Resultados	43
6.1. Modelo <i>Logit</i> com capacidade preditiva entre um e cinco anos antes do evento de insolvência	43
6.2. Modelo Final de Previsão de Insolvência	46
6.3. Teste de Robustez	49
7. Considerações Finais	51
Referências	54
Apêndice	59

Índice de Tabelas

Tabela 1 – Estatística Descritiva da Amostra Total.....	26
Tabela 2 - Matriz de Correlação de Pearson	28
Tabela 3 – Testes às Médias e Medianas	30
Tabela 4 – Classificação das Empresas – Confusion Table	32
Tabela 5 – Erros Tipo I e II derivados da Classificação das Empresas	32
Tabela 6 – Análise dos Coeficientes da Função Discriminante	33
Tabela 7 - Análise de Correlação de Pearson (Structure Matrix)	35
Tabela 8 – Análise de Variância (ANOVA)	36
Tabela 9 - Resultados Antes do Processo de Estandarização das Variáveis	39
Tabela 10 - Resultados Após o Processo de Estandarização das Variáveis.....	40
Tabela 11 – Resultados da Estimação do Modelo Logit 1, 2, 3, 4 e 5 Anos Antes da Insolvência	44
Tabela 12 – Resultados da Estimação do Modelo Logit com Controlo para Efeitos Temporais e Indústria.....	45
Tabela 13 - Resultados da Estimação do Modelo Logit Final	47
Tabela 14 – Resultados da Estimação do Teste Out-of-Sample	49

Índice de Gráficos

Gráfico 1 – Coeficientes Standardized	34
Gráfico 2 – ROC (Receiver Operating Characteristic Curve)	38
Gráfico 3 – Significância dos Coeficientes Estandarizados	41
Gráfico 4 - Curva ROC do Modelo Logit Final	48
Gráfico 5 – Curva ROC do Teste Out-of-Sample	50

Lista de Siglas e Abreviaturas

ADM – Análise Discriminante Múltipla

ANOVA – *Analysis of Variance*

CAE – Classificação das Atividades Económica

EBIT – *Earnings Before Interest and Taxes*

EBITDA – *Earnings Before Interest, Taxes, Depreciation and Amortization*

EUA – Estados Unidos da América

RNA – Redes Neurais Artificiais

ROA – Rentabilidade Operacional do Ativo

ROC – *Receiver Operating Characteristic*

UE – União Europeia

Valor de Mercado – VM

1. Introdução

Ao longo dos tempos, o papel do tecido empresarial na sociedade e na economia tem-se revelado indispensável. Tal deve-se ao facto de as empresas serem órgãos responsáveis pela criação de emprego, pela produção de bens e serviços que vão satisfazer as necessidades dos consumidores e, conseqüentemente, pela criação de lucro e riqueza através de uma gestão de recursos eficiente, o que contribui para o crescimento económico. A riqueza obtida a partir da atividade das empresas é, posteriormente, distribuída pelos proprietários, funcionários, fornecedores, Estado e pela sociedade em geral, revelando o elevado interesse que os *stakeholders* demonstram perante o risco de insolvência de uma organização.

Desde meados do Século XX que se assiste a uma evolução crescente no que diz respeito à internacionalização da economia e à globalização, o que implica uma adaptação das empresas face às mudanças verificadas, sendo estas deparadas frequentemente com novas formas de concorrência e organização. Com o objetivo de ser uma verdadeira união económica com livre circulação de pessoas e bens, a União Europeia (UE) surge em 1992, mas só no início de 1999 é que o Euro nasce formalmente, sendo introduzidas as moedas e notas em circulação a meados de 2002, contando com 19 países¹. Segundo Ferreira & Caetano (2013), a adoção desta moeda única advém vantagens como a redução dos custos transacionais suportados pelos agentes económicos, com a maior eficiência nos investimentos realizados e, em consequência, com a promoção das condições que despoletam as fontes do crescimento económico. Porém, verifica-se um maior risco de contágio entre países, no caso da ocorrência de crises económicas.

No entanto, ao adotar medidas eficazes e eficientes de gestão, as empresas alocam melhor os seus recursos, diminuindo a probabilidade de enfrentarem problemas financeiros e evitam o seu encerramento efetivo ou declaração de insolvência.

Apesar de o mundo empresarial ser influenciado por eventos económicos e financeiros mundiais, considerando a dependência notável entre os países da UE, é essencial recorrer a modelos de previsão de insolvência de forma a prever qual a probabilidade de uma empresa entrar em estado de insolvência e, conseqüentemente, tomar medidas atempadas na tentativa de evitar tal acontecimento.

¹ Países que fazem parte da Zona Euro atualmente: Alemanha, Áustria, Bélgica, Chipre, Eslováquia, Eslovénia, Espanha, Estónia, Finlândia, França, Grécia, Irlanda, Itália, Letónia, Lituânia, Luxemburgo, Malta, Países Baixos e Portugal.

A ocorrência de crises económicas afeta rigorosamente a vida da sociedade em geral. A crise da Grande Depressão, em 1929, desencadeada com a quebra da Bolsa de Nova Iorque, foi um desses exemplos, culminando numa grande recessão económica e num consequente aumento da pobreza. Mais tarde, em 2008, vivenciou-se outra grande crise económica e financeira, a crise do *subprime*, despoletada nos Estados Unidos da América (EUA), com a falência do banco de investimentos, *Lehman Brothers*, alastrando-se ao resto do Mundo, consequência da globalização mundial, sendo a Europa um dos continentes que mais sofreu com este evento. Neste âmbito, inúmeras empresas foram afetadas, passando por sérias dificuldades financeiras, acabando muitas por não conseguir enfrentar tais adversidades, declarando estado de insolvência.

Atualmente, para além da recuperação da crise gerada pela pandemia da Covid-19, em que muitas empresas suspenderam a sua atividade durante um vasto período em prol de evitar o contágio do vírus, estamos a vivenciar um aumento dos preços dos bens e serviços, consequência do conflito entre a Rússia e a Ucrânia. Os efeitos a longo prazo de ambos os acontecimentos poderão gerar uma recessão económica severa, com repercussões superiores à crise *Subprime* de 2008.

Neste contexto, emerge a importância do estudo de modelos de previsão de insolvência, cujo objetivo primordial é criar modelos baseados na análise de rácios económicos e financeiros, com um elevado poder explicativo da situação de insolvência. Inicialmente, os modelos desenvolvidos centravam-se na análise dos rácios de forma isolada, porém, estudos mais recentes evidenciam que, para obter um maior poder explicativo das causas que levam à insolvência, os indicadores financeiros não devem ser analisados de forma isolada. Tal deriva de não serem considerados fatores como, por exemplo, a dimensão das empresas e o seu impacto na previsão de insolvência, bem como variáveis de natureza macroeconómica, dado que em ciclos económicos recessivos as empresas estão mais propensas a enfrentar dificuldades financeiras do que quando estão perante um crescimento económico.

Durante largas décadas, inúmeros estudos foram realizados com o propósito de avaliar e aprimorar estes modelos, destacando-se diversos autores, entre os quais os estudos que servem de base para esta investigação. O estudo pioneiro de Beaver (1966) foca-se num Modelo Univariado, ou seja, na análise isolada de variáveis que apresentam uma relação mais forte com a variável de interesse, i.e., o evento de falência. No entanto, não é averiguada a interação existente entre as variáveis independentes, o que limita a consistência dos resultados obtidos, podendo esta

análise ser complementada pela Análise Discriminante Multivariada (ADM), proposta por Altman (1968), que considera várias variáveis explicativas da situação de falência analisadas em simultâneo; o modelo *Z-score* de Altman (1968) atribui diferentes pesos (*scores*) a cada variável independente em função da sua importância na previsão da situação de falência.

Posteriormente, é apresentado por Ohlson (1980) um modelo baseado na Análise Multivariada, mas que apresenta uma maior eficiência do que a ADM, o Modelo *Logit*. Este modelo estima, através de uma regressão logística, a relação entre a variável dependente, de natureza binária, e as variáveis independentes, contínuas ou categóricas, tornando assim os resultados obtidos mais robustos e precisos face às técnicas propostas por Beaver (1966) e Altman (1968).

Considerando que na Zona Euro se verifica uma uniformidade da política monetária e da legislação no seio das empresas cotadas da UE (Loureiro & Silva, 2021), a aplicabilidade destes modelos às empresas cotadas da Zona Euro demonstra relevância, sendo objeto de estudo nesta análise.

Neste seguimento, o propósito desta dissertação é identificar quais os determinantes com maior capacidade preditiva da situação de insolvência no caso das empresas cotadas da Zona Euro e, desta forma, construir um modelo que revele elevada capacidade da situação de insolvência deste grupo de empresas. Além disso, o modelo desenvolvido deve ser capaz de antecipar este evento até cinco anos antes. Para a sua elaboração foi utilizada uma amostra que engloba o período temporal de 2006 até 2021, inclusive. A informação referente ao *status* das empresas foi retirada da base de dados *ORBIS Europe* da *Moody's Analytics*, enquanto os restantes dados económicos e financeiros foram obtidos a partir da base de dados *Refinitiv DataStream* e *Worldscope*, sendo excluídas da amostra final todas as empresas que operam nos setores altamente regulados, como o financeiro e o energético.

Os dados da amostra foram analisados com recurso à Análise Univariada e à Análise Multivariada, de onde se destacam a Análise Discriminante Múltipla e a Análise Logística para se selecionarem as variáveis com maior capacidade de preverem (em conjunto) uma situação de insolvência. Em concordância com os resultados obtidos nas várias etapas realizadas, o modelo final inclui as seguintes variáveis: Rentabilidade das Vendas, Valor Mercado das Vendas, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão. Este modelo apresenta uma capacidade preditiva de, aproximadamente, 80% até cinco anos antes da declaração de insolvência.

A estrutura desta investigação encontra-se dividida em capítulos, nomeadamente: o Capítulo 2, em que é apresentada a revisão de literatura, com ênfase nos conceitos teóricos capazes de suportar o estudo realizado; o Capítulo 3, onde é indicada a metodologia aplicada na análise empírica; o Capítulo 4, que descreve a amostra e as variáveis, bem como a Análise Univariada; o Capítulo 5, que reporta as etapas da Análise Multivariada, metodologia aplicada para a seleção das variáveis e posterior elaboração do modelo final; o Capítulo 6, onde se apresenta o modelo final e a respetiva análise, bem como uma breve reflexão e discussão dos resultados obtidos; por fim, o Capítulo 7 que diz respeito às considerações finais sobre este estudo e sugestões a incorporar em estudos futuros.

2. Revisão da Literatura

A principal função social de uma empresa é gerir com eficiência os seus recursos para assim obter lucros que são, posteriormente, distribuídos pelas partes interessadas da organização. Independentemente do tipo de setor ou atividade, uma empresa tem de ser capaz de cumprir com as suas funções (Schommer, 2000). No que concerne às empresas cotadas, estas podem estar entre as entidades económicas que mais contribuem para o crescimento económico, isto se os empresários se focarem em inovar e promover o desenvolvimento económico (Robinson, 1998).

Segundo autores como Beck (2000) e Held *et al.* (1999), a globalização conduziu a uma crescente interdependência entre os agentes económicos e sociais, o que se refletiu num aumento das oportunidades de desenvolvimento. Contudo, os mesmos autores alertam para um acréscimo implícito dos riscos enfrentados, bem como uma intensificação da concorrência e competitividade.

Assim, perante as mudanças encaradas ao longo dos anos, é reforçada a necessidade de as empresas adaptarem as suas funções, o que envolve escolhas contínuas e decisões para se ajustarem às oscilações do ambiente em que estão inseridas (Galbraith, 1977). Dentro deste ambiente são abrangidos tanto fatores subjacentes à própria empresa, como outras organizações ou grupos, que acabam por depender desse ambiente e que determinam, de certa forma, as suas características. De acordo com Banner & Gagné (1995), gerir a relação existente entre as organizações e o ambiente envolvente traduz a própria sobrevivência das organizações, através do seu desempenho. Para Alter & Hage (1993), o ramo dos estudos organizacionais não se deve reger pela sua frente tradicional, deve, por outro lado, adotar uma conceção baseada na cooperação e entreajuda, sendo estes o elo principal para compreender a natureza da ordem social. Neste sentido, salienta-se o facto de a dependência mútua ou interdependência entre as empresas, mencionada anteriormente, servir de base para a cooperação.

Uma empresa em atividade é sempre portadora de uma probabilidade de entrar em insolvência. Neste contexto, ao abordar a temática da insolvência das empresas, é essencial ter em consideração que, os termos insolvência e falência, apesar de serem comumente confundidos, não representam a mesma situação a nível económico e financeiro para as empresas. Em acréscimo, o conceito de insolvência não é consensual entre a literatura financeira, como será verificado posteriormente.

Para Beaver (1966), o estado de falência corresponde à suspensão de pagamentos, à existência de contas bancárias a descoberto, ao não pagamento de dividendos ou à insolvência jurídica, ou seja, quando se verifica a incapacidade de uma empresa pagar as suas obrigações financeiras à data do seu vencimento. Altman (1968) considera que o termo falência diz respeito às organizações que estão legalmente falidas, em estado de liquidação ou a passar por um processo de reestruturação, mediante a legislação dos EUA. Segundo Blum (1974), a falência de uma empresa advém da incapacidade de esta pagar as suas dívidas, acabando por recorrer a negociações para amortizar as mesmas ou entrando num processo de insolvência. Por outro lado, Ohlson (1980) optou por adotar uma definição legal quanto ao conceito de falência, isto é, considera falidas as empresas declaradas insolventes e aquelas que entraram num processo legal de insolvência. Por sua vez, Neves (1997) define o termo falência como o declínio da situação económica da organização, que segue um processo constituído por sete fases: na primeira, são recolhidos indícios que têm impacto negativo no desempenho económico da empresa; a segunda fase corresponde àquela em que a empresa apresenta pouca ou nenhuma liquidez; seguindo-se a fase em que a entidade revela incapacidade de cumprir as suas obrigações; posteriormente, verifica-se um fundo de maneio negativo que leva ao acréscimo dos incumprimentos; a quinta fase resulta na insolvência parcial derivada desses incumprimentos; na sexta fase, a empresa entra em falência total, uma vez que o passivo total excede o ativo total; e, por fim, na sétima fase, a empresa enfrenta o processo de liquidar os seus ativos para indemnizar a entidade credora, declarando falência.

Nos estudos de Zmijewski (1984) e Casey & Bartczak (1985), os autores incluíram no grupo de empresas insolventes as que solicitaram insolvência via judicial, definição consistente com outros autores, tais como Diakomihalis (2012) e Lin (2014). Para além de incluírem as empresas que tivessem declarado insolvência ou tivessem sido declaradas como insolventes, Gentry, Newbold, & Whitford (1985) englobam no grupo de empresas insolventes aquelas que foram liquidadas. O estado de insolvência apesar de não exigir o encerramento da empresa, acarreta problemas financeiros, que podem ser apenas temporários devido ao desempenho negativo da mesma (Altman, 1993). Altman & Hotchkiss (1993) referem que o estado de falência não envolve a extinção da empresa, visto que esta pode, durante um período, cumprir as suas obrigações, no entanto, quando deixa de ser capaz de liquidar as suas responsabilidades legais, a empresa entra num estado de “falência legal” e pode ser extinta. Estes autores consideram que insolvência se trata de uma situação em que uma organização não tem liquidez para cumprir com

os seus compromissos, ou seja, o valor do passivo total é superior ao valor do ativo total. Mais tarde, (Antunes, Ribeiro, & Pereira, 2017) consideram que falência ocorre quando uma entidade não é capaz de cumprir as suas obrigações financeiras.

Antão, Peres & Marques (2008) reforçam a ideia de que o conceito de falência ou insolvência não é consensual, visto que se o mesmo for definida do ponto de vista jurídico, então irá variar perante os diferentes países do mundo. Segundo estes, todas as definições, apesar de apresentarem diferenças, revelam características em comum, podendo este conceito ser genericamente definido como uma situação onde uma empresa não tem capacidade para cumprir os seus compromissos financeiros. Assim, Sousa & Oliveira (2014) assumem o termo insolvência no caso de as empresas recorrerem por via judicial à sua liquidação, enquanto falência corresponde à incapacidade de uma empresa fazer face aos seus compromissos financeiros, com recurso aos bens que tem a seu dispor. Neste âmbito, uma empresa declarar insolvência é indicador de uma condição crónica, deixando de ser algo temporário, o que acaba por ser crítico em contexto empresarial.

A necessidade de recorrer a meios que auxiliassem na previsão da probabilidade de ocorrência de situações de insolvência financeira teve início com a crise da Grande Depressão de 1929. Consequentemente, a importância da análise de rácios financeiros tornou-se evidente, verificando-se o declínio dos mesmos, como os rácios de rentabilidade e de liquidez, quando as empresas se aproximavam de uma situação de insolvência financeira (Smith & Winakor, 1935). A crise do *Subprime* que se começou a sentir a partir de meados de 2007, com início nos EUA e repercussões a nível mundial, impulsionou novamente a preocupação de todas as entidades envolvidas no mercado relativamente à questão da solvabilidade (Silva, 2015). Tal como já foi mencionado, há uma clara dependência entre as empresas não só a nível nacional, mas mundial, daí a dimensão e proporções que crises, como as referidas, tomam (Ferreira, 2014). Adicionalmente, Rodrigues Junior, Silva, & Hein (2012) consideram que nenhuma empresa passa de repente de um período de ascensão para o declínio, sendo este um processo gradual, o que torna possível a sua previsão. Altman, Baidya, & Dias (1979), destacam ainda o facto de que todas as empresas estão sujeitas a enfrentar períodos de instabilidade financeira, o que implica que uma empresa poderá deparar-se, a qualquer momento, com problemas financeiros que podem culminar na sua descontinuidade.

Perante isto, é evidenciada a necessidade de analisar e abordar a problemática da insolvência de empresas, isto é, a probabilidade com que estes eventos ocorrem e os seus principais determinantes. Esta matéria demonstra importância no campo da saúde financeira de uma empresa, auxiliando os agentes económicos a prever a evolução futura das empresas e a minimizar o risco de insolvência. Há vários estudos empíricos cujo objetivo final é a criação de modelos eficazes na previsão de falência e insolvência. Estes estudos recorrem, em grande parte, à utilização de rácios económico-financeiros, dado o respetivo elevado grau de confiança na identificação prévia de possibilidade de dificuldades financeiras num futuro próximo (Silva, 2010). Rodrigues Junior, Silva, & Hein (2012) referem que o ponto de partida para obter uma boa precisão passa por identificar, entre as empresas insolventes, características comuns que possibilitem perceber qual o ponto de viragem que desencadeia desequilíbrios financeiros.

As primeiras investigações no ramo da previsão da insolvência surgiram nos anos 30, com a análise de rácios de previsão de insolvência por parte de FitzPatrick (1932), Smith & Winakor (1935), Merwin (1942) e Tamari (1966). Décadas mais tarde, as metodologias adotadas por estes autores foram aprimoradas e surgiram novas metodologias de suporte aos modelos de previsão de insolvência, sendo que, no geral, de acordo com a literatura financeira, a natureza dos respetivos modelos classifica-se como: Modelos Univariados, como Beaver (1966), e Modelos Multivariados, onde se destacam Altman (1968) e Ohlson (1980).

Relativamente ao Modelo Univariado de Beaver (1966), o seu objetivo foi desenvolver uma investigação focada no estudo de rácios financeiros com o propósito de prever riscos de falência para as empresas do setor industrial dos EUA, dado que naquela altura não existia um método eficiente de análise de risco, fazendo com que a concessão de crédito por parte da banca fosse demasiado conservadora, sendo, por vezes, o crédito atribuído insuficiente para alcançar os objetivos pretendidos. Para tal, Beaver (1966) recorreu à utilização de vários rácios financeiros a fim de determinar aqueles que apresentavam maior capacidade preditiva, constatando diferenças, entre as empresas insolventes e as saudáveis, que se tornaram mais evidentes à medida que o ano de insolvência se aproximava. O autor utilizou 30 rácios, repartidos em cinco categorias, nomeadamente liquidez, *cash-flow*, rotação do ativo, endividamento e rentabilidade, para um horizonte temporal de onze anos e amostra de 158 empresas (composta por 79 empresas insolventes e 79 em atividade). Estes rácios foram utilizados para ordenar as empresas por ordem crescente e analisar os dados de acordo com o ponto crítico ou ponto de *cutoff* para cada um dos mesmos. No final deste processo, cada uma das empresas da amostra foi classificada como

insolvente ou não insolvente, consoante o valor do rácio se encontre acima ou abaixo do ponto de *cutoff*. Dos rácios que analisou, nomeadamente $\frac{Cash\ Flow}{Total\ Debt}$, $\frac{Net\ Income}{Total\ Assets}$, $\frac{Total\ Debt}{Total\ Assets}$, $\frac{Working\ Capital}{Total\ Assets}$, $\frac{Current\ Assets}{Current\ Liabilities}$ e *No – credit interval*, o que apresentou melhores resultados preditivos foi o rácio $\frac{Cash\ Flow}{Total\ Debt}$, cuja capacidade de explicar a insolvência empresarial rondava os 87% para um ano antes da falência e 78% para cinco anos antes do evento. As vantagens deste modelo são a facilidade de retirar conclusões relativamente à situação financeira de uma empresa, visto que apenas é necessário ter acesso à sua informação financeira, assim como o facto de permitir comparar amostras grandes, com informação associada a diferentes entidades, de forma a determinar a sua *performance* ao longo do tempo (Moreira, 1998). No entanto, este modelo apresenta uma grande limitação, que se baseia na análise de cada rácio ser realizada de modo individual, não permitindo estudar a relação existente entre os rácios, o que limita a sua eficiência. Para além disso, os dados financeiros são meramente quantitativos, pelo que, a vertente qualitativa, que pode ter influência sobre as organizações, não é analisada. Assim, a comparação das empresas pode ser afetada, uma vez que cada empresa pode apresentar práticas distintas e os indicadores financeiros que, posteriormente, são analisados estão sujeitos às normas pelas quais cada empresa se rege (Neves, 2012). O autor do modelo reconhece ainda que há a possibilidade de se criarem modelos com a combinação de vários rácios, capazes de obter resultados mais robustos, como o caso da Análise Multivariada.

Baseada na Análise Multivariada, surge a Análise Discriminante Múltipla (ADM) desenvolvida por Altman (1968), que propõe o modelo *Z-score*. Por sua vez, este modelo combina vários rácios financeiros e considera a sua evolução conjunta, o que possibilita a deteção de potenciais manipulações da informação contabilística, evitando enviesamentos nos resultados. Assim, foram testados 22 indicadores a partir de uma amostra de 66 empresas industriais cotadas, em que metade eram insolventes e as restantes saudáveis, entre 1946 e 1965. É também de salientar que, além de o período temporal dos dados recolhidos ser o mesmo para ambos os grupos, as empresas saudáveis que integravam a amostra apresentavam semelhanças ao nível do setor em que operam e da sua dimensão, relativamente às empresas insolventes.

A função discriminante proposta por Altman (1968) foi a seguinte²:

$$Z = 0,012X_1 + 0,014X_2 + 0,033X_3 + 0,006X_4 + 0,999X_5$$

Mediante os resultados obtidos com a função *Z-score*, quando $Z > 2,99$, a probabilidade de falência é reduzida; verifica-se incerteza quanto à possibilidade de a empresa entrar em insolvência quando $1,8 < Z < 2,99$; e, se $Z < 1,80$, a probabilidade de falência é elevada.

Com base na ADM, Altman (1968) concluiu que apenas cinco rácios eram significativos para o modelo, isto é, apenas esses rácios evidenciavam diferenças significativas entre empresas insolventes e não insolventes, nomeadamente os rácios de liquidez, rentabilidade, endividamento, solvabilidade e capacidade operacional. Neste âmbito, o rácio com maior significância, representando aquele que melhor explicava a insolvência das empresas, era o $\frac{\text{Working Capital}}{\text{Total Assets}}$, que avalia o nível de liquidez de uma empresa. Em relação à capacidade preditiva deste modelo, a mesma ronda os 95% para um ano antes da insolvência, o que revela um elevado nível de eficiência do modelo. Contudo, nos anos subsequentes, 2, 3, 4 e 5 anos, essa capacidade preditiva diminuiu para 72%, 48%, 29% e 36%, respetivamente. Perante estes resultados, é evidenciado que a deterioração dos rácios é cada vez mais notória com a aproximação da data da falência, conforme foi apurado na Análise Univariada, e que as mudanças mais preocupantes na maioria dos rácios analisados ocorriam entre o terceiro e segundo ano antes da falência.

Em suma, a ADM é uma técnica utilizada para classificar qualquer observação num ou mais grupos dependendo *a priori* das características individuais de cada uma, sendo que classifica a sua variável dependente qualitativamente, isto é, se a mesma é insolvente ou não. Esta técnica estatística serviu de base para estudos subsequentes de diversos autores, entre os quais Deakin (1972), Blum (1974) e Keasey & Watson (1987).

Por um lado, a ADM permite introduzir múltiplos rácios financeiros em simultâneo nos modelos, assumindo também capacidade de combinar estas variáveis independentes (Sung, Chang, & Lee, 1999). Adicionalmente, esta técnica destaca-se pela sua facilidade de aplicabilidade e pela capacidade de extrair resultados com qualidade nos estudos onde foi aplicada (Marín, 1977; Zeytinoglu & Akarım, 2013).

² Sendo que: $X_1 = \frac{\text{Working Capital}}{\text{Total Assets}}$; $X_2 = \frac{\text{Retained Earnings}}{\text{Total Assets}}$; $X_3 = \frac{\text{EBIT}}{\text{Total Assets}}$; $X_4 = \frac{\text{Market Value Equity}}{\text{Book Value Total Debts}}$; $X_5 = \frac{\text{Sales}}{\text{Total Assets}}$.

Por outro lado, apesar da boa capacidade preditiva, vários autores, como por exemplo Eisenbeis (1977) e Pinches (1980), colocaram em causa a eficácia deste modelo para uma obtenção de resultados precisos da previsão de insolvência, baseando-se nas limitações desta análise, como por exemplo a presença de falhas relacionadas com a distribuição das variáveis, a dispersão da matriz de covariâncias, o cálculo individual da significância de cada variável e consequentes erros na classificação dos resultados. Neste seguimento, Sheppard (1994) menciona que a ADM assume que as variáveis explicativas apresentam uma distribuição normal multivariada com diferentes médias, mas matrizes de dispersão iguais, o que implica que, no caso de nenhuma das variáveis apresentar uma distribuição normal, os métodos usados podem resultar numa seleção inapropriada do conjunto dos preditores. Ohlson (1980) acrescenta ainda que o resultado obtido com base na ADM não é fácil de interpretar, considerando que esta análise apenas faz a discriminação das empresas entre os dois grupos – insolventes e não insolventes – e não ilustra a probabilidade de uma empresa entrar em insolvência. Outra crítica do autor deriva de a comparação entre as empresas insolventes e saudáveis ser realizada através de critérios considerados arbitrários, como por exemplo, o setor de atividade e a dimensão.

De forma a ultrapassar as lacunas dos modelos acima referidos, Ohlson (1980) desenvolveu uma nova abordagem para a previsão da insolvência, o Modelo Multivariado de Probabilidade Condicionada, que corresponde a um método econométrico, baseado na Análise Logística, que considera as propriedades da distribuição das variáveis. O Modelo *Logit* representa uma função de probabilidade acumulada aplicada à previsão da insolvência, sendo que o obtido com estimação apresenta dois resultados possíveis – insolvente ou saudável.

Contrariamente à ADM, com este modelo é possível abordar amostras desproporcionais e aumentar a validade da inferência estatística. No entanto, a Análise *Logit* apresenta pressupostos como a ausência de multicolinearidade entre as variáveis explicativas; os erros seguem uma distribuição binomial e são independentes; o contributo das variáveis explicativas é constante e apresentam uma relação de independência com as restantes variáveis explicativas. Segundo Altman & Sabato (2013), este método apresenta vantagens em relação à ADM, tais como a possibilidade de utilização de informação qualitativa, a utilização de uma amostra desproporcional e a possibilidade de afastar a condição de normalidade na distribuição das variáveis.

Para a elaboração do modelo, Ohlson (1980) recolheu dados para um período de seis anos, em que das 2163 empresas sob análise, 2058 eram saudáveis e 105 eram insolventes,

excluindo empresas pequenas e não cotadas e também empresas afetas a serviços públicos, empresas prestadoras de serviços financeiros, companhias de transporte, entre outras pertencentes a setores altamente regulados. Após a recolha da amostra e dos indicadores financeiros, tal como já referido, Ohlson (1980) utilizou um Modelo *Logit* estimado com recurso ao Método da Máxima soma Verosimilhança, um algoritmo capaz de estimar os coeficientes β do modelo que maximizam o logaritmo natural da função de verosimilhança, sendo que, para esta estimação, o autor assumiu que os erros de uma classificação incorreta são iguais para ambos os grupos de empresas, insolventes e saudáveis. Desta forma, definiu um ponto de *cutoff* de 0,5, onde uma empresa seria classificada como insolvente se $P(X_i, \beta) > 0,5$ e saudável quando $P(X_i, \beta) < 0,5$. No entanto, o autor ao longo dos seus estudos procurou um valor que minimizasse a soma dos erros chegando a um ponto de *cutoff* de 0,038. Posto isto, das nove variáveis analisadas, as que se destacaram pela sua capacidade preditiva foram: $SIZE = \log \frac{Total\ Assets}{GDP\ Index\ Price}$, que corresponde à dimensão da empresa; $TLTA = \frac{Total\ Liabilities}{Total\ Assets}$, que representa a estrutura financeira; $NITA = \frac{Net\ Income}{Total\ Assets}$, referente à *performance* da empresa e $CLCA = \frac{Current\ Liabilities}{Current\ Assets}$, que corresponde ao rácio da liquidez geral, para um e dois anos antes da insolvência. Com esta abordagem, o modelo de Ohlson (1980) conseguiu obter uma percentagem de precisão de 96,12% para um ano antes da falência, 95,55% para dois anos antes e 92,84% para três anos antes.

Ohlson (1980) conclui que os resultados obtidos na ADM não eram tão intuitivos como os obtidos na regressão logística, dado que a última técnica torna possível determinar a probabilidade de uma empresa falir. Apesar disso, o Modelo *Logit* revela percentagens de erro superiores à ADM, justificados pelo facto de, em estudos prévios, não ter havido a preocupação de recolher dados de demonstrações financeiras publicadas antes da data de insolvência. Adicionalmente, o autor conclui que o poder de previsão do modelo depende da altura em que a informação é obtida no que concerne à data da insolvência e que uma melhoria significativa do modelo requeria variáveis explicativas adicionais.

Lo (1986) defende que os parâmetros estimados pelo Modelo *Logit* são mais robustos quando comparados com os estimados a partir do modelo baseado na ADM. Tal como referido, nesta metodologia, a variável dependente consiste numa variável dicotómica que adota os valores 1 ou 0, que no caso da estimação de modelos de previsão de falência, correspondem, respetivamente, a empresas insolventes e a empresas que não entraram em estado de insolvência,

sendo consideradas empresas saudáveis. Desta forma, no modelo de regressão logística, a relação entre a probabilidade de falência de uma empresa (P) e o valor dos rácios económico-financeiros dessa empresa em determinado ano é definida pela seguinte expressão: $P(Y_i = 1) = \frac{1}{1+e^{-Z_i}}$, em que Z_i corresponde ao resultado da regressão logística $Z = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_m X_m$. Através do cálculo do índice Z, é possível obter uma determinada probabilidade de falência (P).

Apesar de evidenciar uma melhor capacidade preditiva face às metodologias existentes na época, a investigação de Ohlson (1980) foi alvo de diversas críticas, nomeadamente por Zavgren (1985). Para a sua investigação, Zavgren (1985) recorreu à análise de uma amostra de 90 empresas industriais sediadas nos EUA, das quais 45 eram insolventes e as restantes saudáveis, com dados financeiros para o período compreendido entre 1972 e 1978. Em seguida, estimou modelos *Logit* para cada um dos cinco anos anteriores à falência com base nos seguintes rácios: $\frac{Total\ Income}{Total\ Equity}$, $\frac{Sales}{Net\ PLant}$, $\frac{Inventory}{Sales}$, $\frac{Debt}{Total\ Equity}$, $\frac{Receivables}{Inventory}$, $\frac{Quick\ Assets}{Current\ Liabilities}$, $\frac{Cash}{Total\ Assets}$. Além disso, o autor realizou um teste *Out-of-Sample* de modo a testar o ajustamento dos modelos por si estimados, através de uma amostra de 32 empresas (sendo metade falidas e a outra metade não falidas) para um período compreendido entre 1979 e 1980. A taxa de erro associada a esta amostra de controlo, para um ano antes da ocorrência de falência, foi de 31%. Neste âmbito, Zavgren (1985) concluiu que os rácios de eficiência são mais explicativos a longo prazo e que os rácios de rentabilidade não relevam peso significativo como medidas discriminatórias entre empresas falidas e não falidas. Autores como Platt & Platt (1990) apontam para o facto de o período temporal em que os modelos são estimados e aplicados ser o mesmo, influencia a *performance* dos mesmos, em comparação com modelos que são estimados e aplicados em períodos diferentes. Estes autores identificam também que o comportamento dos rácios pode ser afetado por fatores como o efeito da indústria e a competitividade do mercado, a inflação, alterações nas taxas de juro, e o facto de as amostras serem constituídas por diferentes setores (amostras heterogéneas). É ainda salientado que o Modelo *Logit* tem vantagens associadas, tais como: não assumir uma relação linear entre as variáveis dependentes e independentes, nem exigir que as variáveis sigam uma distribuição normal; a possibilidade de utilizar informação não financeira, isto é, as variáveis independentes podem ser quantitativas e/ou qualitativas; os resultados são mais robustos quando comparados com outras análises multivariadas; e a variável dependente pode ser interpretada como a probabilidade de uma empresa entrar em insolvência.

Contudo, o modelo apresenta limitações tais como: a obrigatoriedade de os grupos dependentes estarem separados; a probabilidade de insolvência seguir uma distribuição logística que varia entre 0 a 1; e a possibilidade de se verificar multicolinearidade entre variáveis, *outliers* e *missing values*.

Após o estudo pioneiro de Ohlson (1980) seguiram-se outros que aplicaram uma metodologia semelhante. Shumway (2001) recorre ao Modelo *Logit* de múltiplos períodos (MLogit) para estimar a probabilidade de insolvência das empresas, sendo este um modelo discreto capaz de capturar mudanças nas empresas, contrariamente aos modelos desenvolvidos por Altman (1968), Ohlson (1980), Taffler (1983) e Zmijewski (1984), em que a probabilidade de falência não muda ao longo do tempo assim como o *status* das empresas, insolvente ou saudável. Zmijewski (1984) propõe um Modelo *Probit* que pressupõe que a probabilidade de falência segue a função de densidade de probabilidade associada à distribuição normal. Para avaliar a probabilidade de uma empresa entrar em insolvência, o autor recorreu à análise de três variáveis explicativas, nomeadamente: $\frac{Net\ Income}{Total\ Assets}$, $\frac{Total\ Debt}{Total\ Assets}$, $\frac{Current\ Assets}{Current\ Liabilities}$, considerando uma amostra de 840 empresas (40 falidas e 800 não falidas), para um horizonte temporal compreendido entre 1972 e 1978. Os resultados deste estudo sugerem que o desempenho deste modelo não apresenta diferenças significativas relativamente aos resultados obtidos com técnicas de amostragem aleatória. Borooah (2002) argumenta que este modelo apenas difere do *Logit* em termos de magnitude dos coeficientes, isto é, a dificuldade da sua interpretação é maior, daí a sua menor utilização, porém, para dados iguais, as conclusões são semelhantes. Da mesma forma, Pampel (2000) conclui que o Modelo *Probit* não é recomendado quando existem muitos casos assimétricos ou com uma distribuição que não a normal. Assim, apesar de a Análise *Probit* permitir obter resultados próximos do Modelo *Logit*, as limitações que tem associadas conduzem a que seja menos utilizado.

Em estudos mais recentes, Altman *et al.* (2017) analisaram o modelo *Z-score* de Altman (1993) em contexto internacional, a partir de uma amostra de 31 países europeus e 3 não europeus. A amostra final englobava 2 602 563 empresas saudáveis e 38 215 empresas insolventes, incluindo empresas públicas e privadas, financeiras e não financeiras, de todos os setores de atividade, para o período compreendido entre 2002 e 2010. Da análise deste modelo, os autores obtiveram resultados favoráveis, com um nível de precisão de cerca de 90%, à medida que incluíram variáveis macroeconómicas relacionadas com as características dos diferentes países. Além disso, os autores testaram a análise de regressão logística aplicada à análise de

desempenho do modelo *Z-score*, em contexto internacional, revelando resultados positivos tanto da análise logística como da ADM.

Recentemente tem sido abordada uma nova metodologia, os modelos *Machine Learning*. Segundo Hastie, Tibshirani, & Friedman (2009), estes algoritmos são uma boa ferramenta para detetar a probabilidade de insolvência dado que são capazes de lidar com um grande número de covariáveis, sendo também eficientes na deteção de relações não lineares. Dentro desta análise, destacam-se as Redes Neurais Artificiais (RNA) que representam um algoritmo muito poderoso, mas também menos transparente, de processamento de informação que testam simultaneamente várias informações, visando a maximização das iteratividades e cruzamento de dados (Gu, Bryan, & Dacheng, 2020). Associado à aplicabilidade deste método estão vantagens como a ausência dos pressupostos da normalidade, linearidade e independência de dados. O primeiro estudo foi desenvolvido por Dumontier (1996), concluindo que as RNA reproduzem o funcionamento do cérebro, mas de forma simplificada. Adicionalmente, o autor concluiu que as RNA geram melhores resultados do que o Modelo Logístico e que as redes que recorrem ao uso dos valores reais dos indicadores obtiveram um menor desempenho do que aquelas que utilizaram a diferença entre os valores dos indicadores.

Mais tarde destaca-se o estudo de Du Jardin (2010), que demonstrou que um modelo para prever a situação de falência baseado em RNA deve ser desenvolvido por meio de métodos de seleção de variáveis apropriados e não com base no que é recorrente na literatura, o que conduz a resultados consideravelmente mais favoráveis do que os modelos ADM e logísticos, com níveis de precisão superiores a 90%.

A elaboração de um modelo que permita estimar a probabilidade de falência relaciona estas empresas falidas com todas as empresas em atividade, fazendo com que os processos aumentem de complexidade nas relações (Zoričák *et al.*, 2020). Apesar disso, as demonstrações financeiras não refletem o estado atual da situação das empresas, apenas se baseiam nos resultados obtidos em anos anteriores, o que condiciona o estudo destes modelos. De salientar ainda que apesar dos avanços na literatura e da evolução tecnológica, os investigadores baseiam-se em modelos empíricos propostos na literatura anterior, tentando desenvolver modelos cada vez mais capazes de responder à questão da probabilidade de insolvência, tornando imprescindível o desenvolvimento e aprimoramento dos modelos (Zopounidis & Doumpos, 1999). No entanto, segundo autores como Cardoso (1996) e Morris (1997), apesar dos significativos

desenvolvimentos a nível dos modelos de previsão de insolvência, não existe uma estrutura conceptual teórica que explique de forma clara e vasta como e porque é que as empresas entram em falência.

3. Metodologia

Esta investigação será baseada em dois tipos de análise: univariada e multivariada. Primeiramente surge a Análise Univariada, uma técnica que permite avaliar o poder explicativo de cada variável independente face à variável dependente e, deste modo, identificar quais as variáveis explicativas que apresentam uma relação mais forte com a variável explicada. Esta foi uma das primeiras técnicas a ser aplicada a dados financeiros, enaltecendo o trabalho realizado por Beaver (1966), tal como mencionado no capítulo anterior. No entanto, de acordo com a literatura, esta análise apresenta limitações notórias uma vez que os indicadores são analisados de forma individual e, conseqüentemente, a interpretação dos resultados obtidos pode ser inconclusiva no que diz respeito ao desempenho financeiro das empresas (Altman, 1968). Tendo em consideração que a Análise Multivariada é uma análise mais complexa comparativamente à Análise Univariada, será alvo de um maior destaque ao longo deste estudo.

A Análise Multivariada explora as variáveis independentes em simultâneo, considerando que um dos grupos em análise foi alvo de um tipo de tratamento e o outro não, ou seja, para o estudo em causa, num dos grupos foi declarado estado de insolvência, contrariamente ao outro, em que as empresas são saudáveis, sendo este denominado de grupo de controlo. Apesar de uma melhor capacidade explicativa, esta técnica também apresenta limitações, sendo aplicadas ao longo desta dissertação metodologias com o propósito de mitigar possíveis dificuldades associadas a esta análise, como por exemplo, a presença de enviesamentos ou problemas de endogeneidade (Loureiro & Silva, The impact of cross-delisting from U.S. on firms' financial constraints, 2020). De forma a obter o modelo final, dentro da Análise Multivariada serão aplicadas duas técnicas para seleccionar as variáveis que melhor identificam a possibilidade de ocorrência de insolvência, nomeadamente a Análise Discriminante Múltipla (ADM) aplicada no modelo *Z-score* de Altman (1968) e a Análise Logística proposta por Ohlson (1980).

A ADM é uma análise estatística cujo objetivo é obter, através do estudo das diferenças entre dois grupos, uma função que minimize a probabilidade de classificação incorreta dos membros de um grupo. Por outras palavras, a melhor combinação linear de variáveis independentes para discriminar elementos pertencentes a diferentes grupos é aquela que permite minimizar os erros de uma classificação incorreta. Nesta investigação, esta técnica vai permitir classificar a variável dependente qualitativamente, ou seja, quais das empresas em análise se inserem no grupo de insolventes e quais se inserem no grupo de empresas saudáveis.

Adicionalmente, fornecerá também informação acerca das variáveis preditoras sob análise, i.e., quais as mais relevantes para a classificação dos grupos (e que também se caracterizam por gerarem uma menor variação dentro de cada grupo).

Tal como previamente mencionado, a ADM será a primeira metodologia multivariada a ser aplicada. Neste sentido, a função discriminante do modelo a testar é a seguinte:

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_i X_j \quad (1)$$

Em que Y corresponde ao *score* da função, β_i representam os coeficientes da função e X_j as variáveis discriminantes da situação de insolvência. Como primeira etapa da ADM são estabelecidas características que distingam os grupos que estão a ser analisados e, posteriormente, os coeficientes discriminantes são multiplicados pelas respetivas variáveis explicativas de acordo com a função discriminante, resultando num índice que possibilita a classificação de uma empresa como insolvente ou saudável, mediante os grupos definidos *a priori* (Sage, 2015).

Para que a análise revele resultados não enviesados, as variáveis discriminantes têm de apresentar uma distribuição normal, as matrizes de variância e covariância de cada grupo têm de ser aproximadamente iguais e as variáveis devem apresentar diferenças significativas entre os grupos (p.e., Reis, 2001). Ohlson (1980) salienta a existência de *outliers* como uma limitação da ADM por poder afetar negativamente a aplicabilidade deste método, sendo que a validade da inferência estatística pode estar em causa devido às médias e variâncias deturpadas resultantes dessas observações anormais. Contudo, segundo Eisenbeis (1977), os desvios à distribuição normal, pelo menos na área económica e financeira são, na prática, mais recorrentes do que a presença de variáveis com distribuição normal, dificultando a mitigação desta limitação. Empiricamente, rejeitando-se o pressuposto da normalidade das variáveis, o modelo pode incluir um conjunto de preditores inadequados à previsão de falência (Sheppard, 1994). No que diz respeito à igualdade das matrizes de variância e de covariância, esta hipótese não deverá ser rejeitada, uma vez que, normalmente se verifica a inexistência de multicolinearidade entre as variáveis. Além disso, a aplicação deste modelo requer previamente informação acerca do momento que antecede a falência. Contudo, após se formular o modelo baseado na ADM, este método é de fácil aplicabilidade, com boa capacidade para incorporar múltiplos rácios financeiros,

bem como para combinar variáveis independentes, fornecendo, assim, resultados mais consistentes do que a Análise Univariada (Altman, 1968).

Neste sentido, apesar de alguns dos pressupostos da ADM terem de ser relaxados (como é o caso da normalidade das variáveis), esta análise será realizada dado que permite fornecer uma avaliação objetiva das diferenças entre as empresas insolventes e as empresas saudáveis, com base num conjunto de variáveis independentes expressas sob a forma de rácios económico-financeiros.

Posterior a esta análise, tratada como uma análise preliminar, segue-se a Análise Logística. A Análise Logística recorre a um modelo de regressão de estimação logística, sendo definida uma variável dependente dicotómica, que assume o valor de 1 no caso de uma condição se verificar, que neste caso corresponde a uma empresa ser insolvente, e de 0 caso contrário, i.e., englobando as empresas que se encontram em atividade, não enfrentando graves dificuldades financeiras. Ohlson (1980) desenvolveu um modelo de previsão de insolvência baseado nesta metodologia, visando atenuar alguns dos pressupostos relativamente à ADM, previamente mencionados. Neste âmbito, o modelo logístico pretende estimar a probabilidade de determinado evento acontecer em função das variáveis independentes, recorrendo ao estimador da máxima verosimilhança, dando origem à seguinte equação:

$$Prob(Y_i = 1) = \beta'X_{i,t=-1} + \lambda_j + \gamma_t + \varepsilon_{it} \quad (2)$$

Onde, Y_i representa a variável binária, que assume o valor 1 para determinado evento e 0, caso contrário; $\beta'X_{i,t=-1}$ inclui o conjunto de variáveis independentes desfasadas no tempo relativamente ao momento de ocorrência da insolvência; λ_j e γ_t correspondem aos parâmetros que permitem controlar a heterogeneidade não observada em relação à indústria e aos anos que constam no painel, respetivamente e, por fim, o termo ε_{it} diz respeito ao erro da regressão. Neste seguimento, a probabilidade acumulada da função é calculada a partir da seguinte expressão:

$$Prob(Y_i = 1) = \frac{1}{1+e^{-\beta'X_{i,t=-1}}} \quad (3)$$

Esta análise fornece resultados mais precisos relativamente à ADM, pelo que as conclusões baseadas nos seus resultados terão maior peso na seleção das variáveis que irão incorporar o modelo final.

4. Descrição da Amostra e Variáveis

Tal como já mencionado, o objetivo desta dissertação prende-se com a proposta de um modelo empírico que permita prever a insolvência financeira para as empresas cotadas da Zona Euro, baseado no modelo de regressão logística proposto por Ohlson (1980). Este modelo deve apresentar uma capacidade preditiva de assinalar a situação de insolvência até cinco anos antes desse acontecimento, bem como quais os determinantes que influenciam a insolvência das empresas.

De seguida, será descrita a amostra recolhida para este estudo, identificando os dados empíricos e a sua natureza, o horizonte temporal, bem como possíveis limitações que surgiram ao longo do processo de recolha de dados.

4.1. Amostra

Os dados financeiros que compõe a amostra foram recolhidos da base de dados *Refinitiv DataStream* e *Worldscope* e a informação sobre o *status* da empresa – insolvente ou saudável – foi obtido a partir da *ORBIS Europe* da *Moody's Analytics*, sendo que a amostra final agregará dados financeiros e contabilísticos extraídos entre 2006 e 2021, de forma a abranger um período de tempo considerável.

De acordo com a classificação atribuída pela *ORBIS Europe*, as empresas saudáveis são aquelas que desempenham em pleno as suas atividades e que representam situação financeira considerada sustentável, representando 930 empresas durante o período em análise com o *status* “*Active*”. Paralelamente, contabilizam-se 51 empresas consideradas insolventes, cujo *status* é “*Active (insolvency proceedings)*” e “*Banckruptcy*”, sendo estas as empresas que se encontram em processo de insolvência ou em estado de falência técnica (i.e., capitais próprios negativos).

Ambos os grupos de empresas foram sujeitos à aplicação de filtros, nomeadamente no que diz respeito ao setor de atividade económica, sendo que empresas do setor financeiro e energético, classificadas com os Códigos de Atividade Económica (NACE *Code*) 35, 36, 64, 65, 66 e 68 foram excluídas da amostra por serem setores altamente regulados. Além disso, apenas são integradas na amostra empresas cotadas nas Bolsas de Valores Mobiliários da Zona Euro.

4.2. Variáveis

A variável dependente deste modelo logístico corresponde à classificação das empresas como insolventes ou saudáveis, sendo definida por *Insolvency*. Esta variável assume uma natureza binária, isto é, adota o valor de 1 no caso de a empresa ser insolvente e de 0 se a empresa é considerada saudável.

Relativamente às variáveis explicativas, foram selecionados rácios económico-financeiros, que são comumente utilizados no estudo de previsão de insolvência de acordo com vários autores citados no Capítulo 2. Estes rácios refletem a situação financeira das empresas ao nível do seu endividamento, solvabilidade, liquidez e rentabilidade. A sua análise demonstra elevada importância, uma vez que quando se está perante oscilações nos valores que estes assumem, uma empresa pode estar em risco de entrar em dificuldades financeiras que culminem na sua insolvência. Adicionalmente, as variáveis independentes foram sujeitas a um processo de *winsorizing* em 1% em cada cauda, de modo a mitigar o impacto de *outliers* na análise empírica, permitindo que os valores extremos das variáveis igualem o percentil 1 e 99.

Assim, apresentam-se de seguida as variáveis consideradas determinantes da situação de insolvência, expressas sob a forma de rácios económico-financeiros:

Rácio do Fundo de Maneio: o valor deste rácio expressa a diferença entre o ativo corrente e o passivo corrente em relação ao ativo total, ou seja, manifesta a proporção de fundos necessários à liquidação das obrigações da empresa a curto prazo. Valores baixos neste rácio estão associados a riscos mais elevados de insolvência, pelo que o ideal serão rácios com valores superiores a zero (Jackendoff, 1962).

$$\text{Fundo de Maneio} = \frac{\text{Ativo Corrente} - \text{Passivo Corrente}}{\text{Ativo Total}}$$

Rentabilidade Económica: este rácio permite a comparação entre os resultados operacionais, isto é, antes de juros e impostos, e o valor do ativo total. O valor do rácio permite inferir acerca do grau de eficiência da alocação dos ativos pela empresa e, conseqüentemente, se obtém resultados positivos ou negativos. Assim, quanto maior for o rácio, melhor será o desempenho da empresa, conduzindo a uma menor probabilidade de insolvência (Beaver, 1966).

$$\text{Rentabilidade Económica} = \frac{\text{EBIT}}{\text{Ativo Total}}$$

Rentabilidade das Vendas: a análise deste rácio fornece informação acerca da margem líquida que uma empresa gera por cada unidade monetária obtida através das suas vendas. Neste sentido, valores elevados neste rácio traduzem-se numa maior eficiência da empresa (Jagels & Coltman, 2004).

$$\text{Rentabilidade Vendas} = \frac{\text{Resultado Líquido}}{\text{Vendas}}$$

Rentabilidade Operacional do Ativo (ROA): este rácio corresponde à divisão dos resultados antes de juros, impostos, depreciações e amortizações de uma empresa, pelo seu ativo total. Valores superiores neste rácio diminuem a probabilidade de uma empresa entrar em estado insolvente por corresponderem a um desempenho operacional positivo da mesma (Altman *et al.*, 2017).

$$ROA = \frac{EBITDA}{\text{Ativo Total}}$$

Rentabilidade do Ativo: através da análise deste rácio é possível observar a relação entre os resultados líquidos e o ativo total. Quanto maior for o valor do rácio, menor será a probabilidade de uma empresa incorrer em insolvência, uma vez que permite avaliar se os seus ativos são explorados de forma rentável ou não (Moreira & Tavares, 2014).

$$\text{Rentabilidade Ativo} = \frac{\text{Resultado Líquido}}{\text{Ativo Total}}$$

Valor de Mercado (VM) das Vendas: este rácio representa o valor de mercado do capital próprio por cada unidade monetária obtida através das vendas da empresa, que geram receita para a mesma (Damodaran, 2012).

$$VM \text{ Vendas} = \frac{\text{Valor de Mercado do Capital Próprio}}{\text{Vendas}}$$

Margem EBIT: esta margem traduz a margem de lucro operacional, sem contabilizar os juros e impostos, de cada unidade monetária gerada pelas vendas da empresa. Deste modo, quanto mais baixo for o valor deste rácio, em anos consecutivos, maior a probabilidade de uma empresa enfrentar uma potencial situação de insolvência (Altman & Sabato, 2013).

$$\text{Margem EBIT} = \frac{EBIT}{\text{Vendas}}$$

Liquidez Geral: este rácio revela a capacidade de a empresa cumprir com as suas responsabilidades de curto prazo. Valores superiores a 1 evidenciam que não há problemas a nível de liquidez, no entanto, se o valor do rácio for inferior a 1, prevê-se um desequilíbrio financeiro que, mais facilmente, pode culminar num aumento da probabilidade de a empresa se tornar insolvente (Jackendoff, 1962).

$$\text{Liquidez Geral} = \frac{\text{Ativo Corrente}}{\text{Passivo Corrente}}$$

Liquidez Imediata: este rácio calcula a quantidade de ativos de uma empresa que podem ser rapidamente convertidos em dinheiro em relação ao passivo exigível no curto prazo. Analogamente, empresas com margens mais reduzidas de liquidez são, normalmente, mais propensas a enfrentar uma situação de insolvência (Beaver, 1966).

$$\text{Liquidez Imediata} = \frac{\text{Caixa e Equivalentes}}{\text{Passivo Corrente}}$$

Rotação Ativo: fornece informação sobre o grau de eficiência de alocação dos recursos da empresa, através do número de vendas que é gerado em função do investimento em ativos pela mesma. Para uma melhor eficiência e menor risco de insolvência, o valor do rácio deve ser o mais elevado possível (Altman, 1968).

$$\text{Rotação Ativo} = \frac{\text{Vendas}}{\text{Ativo Total}}$$

(Prazo médio) Inventário: este rácio traduz o prazo médio de inventários em armazém, pelo que quantifica o número médio de dias de permanência dos *stocks* em armazém, face ao volume de vendas. Quanto menor o valor do rácio, menor é o tempo médio que o inventário permanece na empresa.

$$\text{Inventário} = \frac{\text{Inventário}}{\text{Vendas}} \times 365$$

Endividamento: através do valor deste rácio é possível aferir a percentagem do ativo financiada pela dívida. Se este rácio apresentar valores elevados, há um maior risco de insolvência (Ohlson, 1980).

$$\text{Endividamento} = \frac{\text{Dívida Total}}{\text{Ativo Total}}$$

Solvabilidade: permite avaliar a alavancagem financeira de uma empresa (Hatfield, Cheng, & Davidson, 1994). Rácios elevados indicam uma menor alavancagem e, conseqüentemente, melhor saúde financeira (Altman, 1968).

$$\text{Solvabilidade} = \frac{\text{Valor de Mercado do Capital Próprio}}{\text{Dívida Total}}$$

Ganhos Retidos: através deste rácio é possível observar a relação entre os ganhos retidos e o ativo total (Altman, 1968). Este rácio reflete os lucros que transitaram do ano n para o ano $n+1$ para financiar a atividade da empresa, estando implícita a idade da mesma. Para o autor, uma empresa jovem provavelmente apresenta menores resultados transitados do que uma empresa mais antiga. Desta forma, algumas empresas poderão ser classificadas de forma errada por serem mais recentes.

$$\text{Ganhos Retidos} = \frac{\text{Ganhos Retidos}}{\text{Ativo Total}}$$

Crescimento das Vendas: representa o crescimento percentual e anual das vendas, entre dois anos consecutivos. Uma taxa de crescimento de vendas elevada apresenta uma menor probabilidade de insolvência implícita.

$$\text{Crescimento Vendas} = \frac{\text{Vendas}_n}{\text{Vendas}_{n-1}} - 1$$

Dimensão: será captada pelo logaritmo do ativo total, que funciona como uma *proxy* da dimensão da empresa (Domenichelli, 2018).

$$\text{Dimensão} = \log \text{Ativo Total}$$

Q(Tobin): rácio que estabelece a relação entre o valor de mercado do capital próprio e os ativos totais, face ao valor contabilístico capital próprio. Este rácio permite avaliar a forma como a empresa está a operar (Zarei, Dadashi, & Akbari, 2012).

$$Q(\text{Tobin}) = \frac{\text{Valor de Mercado do Capital Próprio} + \text{Ativo Total}}{\text{Valor Contabilístico Capital Próprio}}$$

4.3. Estatística Descritiva

Para tratamento estatístico da amostra, bem como os respetivos resultados, recorreu-se à análise dos dados através do *Stata Statistical Software*, versão 16.

A Tabela 1 apresenta a estatística descritiva das variáveis independentes incluídas neste estudo, durante o período de 2006 a 2021, inclusive. Desta forma, os valores da Tabela 1 correspondem aos dados estatísticos das empresas consideradas saudáveis (Painel A) e aos dados do tratamento estatístico das empresas insolventes (Painel B).

Comparando ambos os painéis da Tabela 1, é possível observar que variáveis que se destacam por serem indicadoras de um maior risco de insolvência, como por exemplo, a Rentabilidade Económica, Rentabilidade Vendas, Margem EBIT e Ganhos Retidos, que apresentam valores médios negativos inferiores nas empresas insolventes, face às empresas saudáveis. Além destas, o prazo médio de inventário em armazém correspondente à variável Inventário, VM Vendas, Liquidez Geral e a Solvabilidade apresentam valores médios positivos superiores nas empresas insolventes, em relação às saudáveis, o que demonstra que estas variáveis são reveladoras de incapacidade de exercício da atividade operacional, ou seja, a empresa tem dificuldade em escoar os seus produtos/serviços.

No entanto, variáveis como a Rentabilidade Vendas, Rotação Ativo, Endividamento e Dimensão apresentam valores médios positivos superiores nas empresas saudáveis relativamente às insolventes, sendo assim consideradas variáveis que influenciam positivamente o desempenho das empresas. A presença da variável Endividamento neste grupo pode ser justificada pelo facto de este rácio ser calculado com recurso à dívida financeira, o que pode evidenciar que potenciais entidades credoras (p.e., bancos) já estariam a antecipar um cenário de falência, prevenindo-o.

Da mesma forma, variáveis que possuam valores superiores nas empresas saudáveis comparativamente com as empresas insolventes contribuem para um menor risco de insolvência, como é o caso da variável ROA.

Tabela 1 – Estatística Descritiva da Amostra Total

Painel A – Estatística Descritiva do Grupo de Empresas Saudáveis

Variáveis	N	Média	D.P.	P25	Mediana	P75
Fundo de Maneio	9798	0,1415	0,3180	-0,0075	0,1420	0,3228
Rentabilidade Económica	9672	0,0042	0,2228	-0,0227	0,0416	0,0875
Rentabilidade Vendas	9677	-0,3219	2,0883	-0,0467	0,0176	0,0599
ROA	9564	0,0483	0,2082	0,0230	0,0834	0,1352
Rentabilidade Ativo	9866	-0,0350	0,2285	-0,0431	0,0176	0,0569
VM Vendas	9677	4,0606	21,1029	0,2879	0,6408	1,3723
Margem EBIT	9499	-0,2220	1,6627	-0,0221	0,0413	0,0963
Liquidez Geral	9799	2,3098	3,7593	0,9804	1,3997	2,2089
Liquidez Imediata	9788	0,8576	2,1605	0,0808	0,2359	0,6696
Rotação Ativo	9867	1,0246	0,6840	0,5543	0,9380	1,3458
Inventário	9521	56,8563	84,6098	6,6599	35,0291	72,9370
Endividamento	9801	0,2434	0,2286	0,0490	0,2003	0,3678
Solvabilidade	8714	115,1420	694,8847	0,8637	2,5411	8,6595
Ganhos Retidos	9686	-0,2384	1,3041	-0,1488	0,0549	0,2296
Crescimento Vendas	8269	0,0715	0,4559	-0,0775	0,0282	0,1414
Dimensão	9867	11,3148	1,9629	9,9577	11,2273	12,6164
Q(Tobin)	9840	5,3209	9,6461	2,8246	4,0325	5,9777

Painel B – Estatística Descritiva do Grupo de Empresas Insolventes

Variáveis	N	Média	D.P.	P25	Mediana	P75
Fundo de Maneio	45	-0,0826	0,5657	-0,3885	0,0292	0,2931
Rentabilidade Económica	45	-0,2476	0,4874	-0,3086	-0,0338	0,0184
Rentabilidade Vendas	38	-3,8262	7,0485	-1,1346	-0,0765	0,0074
ROA	43	-0,1867	0,4531	-0,2781	-0,0167	0,0735
Rentabilidade Ativo	45	-0,2878	0,5192	-0,3388	-0,0439	0,0005
VM Vendas	38	58,6828	89,3061	0,3592	0,9386	201,3542
Margem EBIT	38	-3,1247	5,8103	-1,0092	-0,0424	0,0538
Liquidez Geral	45	2,6095	5,0972	0,1559	1,0514	2,4012
Liquidez Imediata	45	0,5933	1,2008	0,0048	0,0646	0,3825
Rotação Ativo	45	0,5657	0,8304	0,0010	0,1209	1,0523
Inventário	38	98,2187	115,0087	25,6369	71,0262	120,2008
Endividamento	45	0,1052	0,1160	0,0000	0,0742	0,1669
Solvabilidade	32	216,2574	1113,7600	1,0735	9,5044	26,5698
Ganhos Retidos	45	-1,4880	2,3828	-1,6786	-0,4849	-0,0695
Crescimento Vendas	32	-0,0360	0,7536	-0,4223	-0,2016	0,1008
Dimensão	45	10,0186	2,0124	8,8870	9,9173	11,0912
Q(Tobin)	45	2,4691	8,2150	1,9388	3,1335	6,3672

Nota: A Tabela 1 apresenta a análise descritiva referente ao grupo de empresas saudáveis (Painel A) e insolventes (Painel B): Número de Observações (N), Média, Desvio Padrão (D.P.), Percentil 25° (P25), Mediana que corresponde ao Percentil 50° e Percentil 75° (P75). As variáveis são as seguintes: Fundo de Maneio, Rentabilidade Económica, Rentabilidade Vendas, ROA, Rentabilidade Ativo, VM Vendas, Margem EBIT, Liquidez Geral, Liquidez Imediata, Rotação Ativo, Inventário, Endividamento, Solvabilidade, Ganhos Retidos, Crescimento Vendas, Dimensão, Q(Tobin), construídas conforme descrito no Subcapítulo 4.2.

4.4. Análise de Correlação de *Pearson*

A Correlação de *Pearson* é uma medida estatística que avalia a relação linear entre duas variáveis quantitativas contínuas, ou seja, permite determinar o grau de correlação entre as variáveis, a partir da análise do coeficiente de correlação, ρ , que varia de -1 a 1, inclusive. Esta análise visa uma interpretação puramente estatística dos resultados e antecede a análise multivariada. Além disso, trata-se de uma análise que apresenta limitações pois os resultados podem ser influenciados pela existência de *outliers* que possam ter resistido ao processo de *winsorizing*.

Da análise deste coeficiente, é possível aferir se as variáveis se correlacionam positiva ou negativamente, ou seja, quando o valor do coeficiente é positivo, está implícita uma correlação positiva e, conseqüentemente, a correlação é negativa se o sinal do coeficiente for negativo. Para além de ser possível a análise da sua inferência estatística, o valor do coeficiente é capaz de indicar a intensidade com que duas variáveis se relacionam, uma vez que quando o valor do coeficiente assume os valores -1 ou 1, estamos perante uma correlação linear perfeita negativa ou positiva, respetivamente, e quando a correlação é 0, estamos perante uma ausência de correlação.

De acordo com os resultados apresentados na Tabela 2, as empresas insolventes estão positivamente (e significativamente) correlacionadas com as variáveis VM Vendas e dias de Inventário em armazém e negativamente (e significativamente) correlacionadas com os rácios de Rentabilidade, Fundo de Maneio, Rotação Ativo, Endividamento, Ganhos Retidos e Dimensão. No entanto, é de destacar que os coeficientes de correlação são considerados de baixa intensidade (são inferiores a 0,30 na escala apresentada por Hinkle, Wiersm, & Jurs, 2003).

É também de destacar que a maior parte das variáveis explicativas são estatisticamente significativas ao nível de, pelo menos, 5%.

Tabela 2 - Matriz de Correlação de *Pearson*

Variáveis	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)
<i>Insolvency</i> (1)	1																
Fundo de Maneio (2)	-0,047*	1															
Rentabilidade Económica (3)	-0,073*	0,296*	1														
Rentabilidade Vendas (4)	-0,102*	0,101*	0,501*	1													
ROA (5)	-0,074*	0,242*	0,971*	0,511*	1												
Rentabilidade Ativo (6)	-0,073*	0,328*	0,969*	0,527*	0,938*	1											
VM Vendas (7)	0,155*	0,006	-0,290*	-0,767*	-0,306*	-0,299*	1										
Margem EBIT (8)	-0,107*	0,100*	0,537*	0,978*	0,541*	0,541*	-0,753*	1									
Liquidez Geral (9)	0,005	0,503*	0,042*	-0,080*	0,016	0,056*	0,189*	-0,084*	1								
Liquidez Imediata (10)	-0,008	0,444*	-0,001	-0,124*	-0,028*	0,012	0,234*	-0,130*	0,878*	1							
Rotação Ativo (11)	-0,045*	0,084*	0,123*	0,217*	0,142*	0,117*	-0,227*	0,208*	-0,147*	-0,182*	1						
Inventário (12)	0,031*	0,064*	-0,117*	-0,177*	-0,152*	-0,109*	0,096*	-0,173*	0,080*	-0,048*	-0,237*	1					
Endividamento (13)	-0,041*	-0,514*	-0,151*	-0,020	-0,143*	-0,198*	-0,082*	-0,009	-0,251*	-0,234*	-0,151*	0,148*	1				
Solvabilidade (14)	0,009	0,149*	-0,029*	-0,227*	-0,034*	-0,045*	0,302*	-0,192*	0,193*	0,247*	-0,039*	-0,020	-0,196*	1			
Ganhos Retidos (15)	-0,065*	0,248*	0,569*	0,392*	0,561*	0,596*	-0,269*	0,387*	-0,020	-0,070*	0,078*	-0,094*	-0,156*	-0,081*	1		
Crescimento Vendas (16)	-0,015	0,069*	0,103*	0,066*	0,094*	0,106*	-0,019	0,064*	-0,043*	-0,020	0,068*	-0,081*	-0,063*	0,025	0,064*	1	
Dimensão (17)	-0,044*	-0,072*	0,289*	0,186*	0,258*	0,285*	-0,153*	0,192*	-0,194*	-0,166*	-0,067*	-0,042*	0,172*	-0,078*	0,341*	0,002	1
Q(Tobin) (18)	-0,020	-0,017	-0,019	-0,056*	-0,017	-0,025*	0,092*	-0,056*	-0,046*	-0,037*	0,036*	-0,017	-0,002	0,017	-0,035*	0,048*	-0,011

Nota: A Tabela 2 reporta os coeficientes de Correlação de *Pearson*. As variáveis utilizadas nesta análise são: Fundo de Maneio (2), Rentabilidade Económica (3), Rentabilidade Vendas (4), ROA (5), Rentabilidade Ativo (6), VM Vendas (7), Margem EBIT (8), Liquidez Geral (9), Liquidez Imediata (10), Rotação Ativo (11), Inventário (12), Endividamento (13), Solvabilidade (14), Ganhos Retidos (15), Crescimento Vendas (16), Dimensão (17), Q(Tobin) (18), construídas conforme descrito detalhadamente no Subcapítulo 4.2.* Significância estatística ao nível de, pelo menos, 5%.

4.5. Análise Univariada

A Análise Univariada recorre à utilização de uma única variável, ou seja, a partir de um teste de classificação dicotômica e da análise isolada de apenas uma variável do conjunto das variáveis selecionadas, é possível classificar as observações da amostra como insolventes ou não insolventes (Cook & Nelson, 1998). O trabalho de Beaver (1966) é uma forte referência neste tipo de análise de previsão de insolvência, sendo esta considerada uma técnica fácil de aplicar e bastante versátil.

Os testes paramétricos testam a significância estatística dos resultados assumindo o pressuposto de que seguem sempre uma determinada distribuição de probabilidade; por outro lado, os testes não paramétricos sustentam a sua análise rejeitando qualquer pressuposto do domínio da distribuição probabilística (Areal & Armada, 1999). Desta forma, com esta análise, pretende-se testar as diferenças às médias e medianas das variáveis em estudo entre as empresas insolventes e saudáveis, recorrendo-se a um teste paramétrico às médias, o *t-test*³, e a um teste não paramétrico às medianas, *Wilcoxon-Mann-Whitney rank sum test*⁴.

A Tabela 3 apresenta os resultados obtidos dos testes paramétricos e não paramétricos.

³ O teste paramétrico às médias, designado *t-test*, é definido pelas hipóteses: H_0 : Diferença das médias das variáveis entre os dois grupos independentes = 0; H_1 : caso contrário.

⁴ O teste não paramétrico às medianas, designado *Wilcoxon-Mann-Whitney rank sum test*, é definido pelas hipóteses: H_0 : Diferença das medianas das variáveis entre os dois grupos independentes = 0; H_1 : caso contrário.

Tabela 3 – Testes às Médias e Medianas

Variáveis	Diferenças entre médias		Diferenças entre medianas	
	P-value		P-value	
Fundo de Maneio	0,0110	**	0,0120	**
Rentabilidade Económica	0,0020	***	0,0000	***
Rentabilidade Vendas	0,0040	***	0,0000	***
ROA	0,0020	***	0,0000	***
Rentabilidade Ativo	0,0020	***	0,0000	***
VM Vendas	0,0010	***	0,0030	***
Margem EBIT	0,0040	***	0,0000	***
Liquidez Geral	0,6960		0,0230	**
Liquidez Imediata	0,1490		0,0000	***
Rotação Ativo	0,0010	***	0,0000	***
Inventário	0,0330	***	0,0020	***
Endividamento	0,0000	***	0,0000	***
Solvabilidade	0,6120		0,0340	**
Ganhos Retidos	0,0010	***	0,0000	***
Crescimento Vendas	0,4260		0,0020	***
Dimensão	0,0000	***	0,0000	***
Q(Tobin)	0,0250	**	0,0250	**

Nota: A Tabela 3 reporta o p-value dos testes às médias (*t-test*) e às medianas (*Wilcoxon-Mann-Whitney rank sum test*) em função da *Insolvency*, isto é, empresas insolventes vs empresas saudáveis. As variáveis são as descritas detalhadamente no Subcapítulo 4.2.

De acordo com os resultados reportados na Tabela 3, quanto à diferença das médias das variáveis, entre os dois grupos independentes, apenas as variáveis Liquidez Geral, Liquidez Imediata, Solvabilidade e Crescimento Vendas não apresentam diferenças estatisticamente significativas. As variáveis Fundo de Maneio e Q(Tobin) são estatisticamente significativas ao nível de 5%. As diferenças nos rácios de Rentabilidade, VM Vendas, Margem EBIT, Rotação Ativo, Inventário, Endividamento, Ganhos Retidos e Dimensão são estatisticamente significativas ao nível de 1%.

Relativamente às medianas, as diferenças entre as variáveis são todas estatisticamente significativas ao nível de, pelo menos, 5%.

Esta análise preliminar destaca todos os rácios de Rentabilidade, a as variáveis VM Vendas, Rotação Ativo, Inventário, Endividamento, Ganhos Retidos e Dimensão como diferenciadoras da situação financeira das empresas da amostra - insolventes *versus* saudáveis.

5. Análise Multivariada

Este método estatístico assume a explicação da variável dependente sob o ponto de vista de uma interação simultânea de múltiplas variáveis e suas combinações. Desta forma, surgem modelos que passam a beneficiar de uma maior consistência na discriminação dos grupos definidos *a priori* e na detecção das relações causa-efeito entre a variável dependente e as variáveis explicativas, nomeadamente, a Análise Discriminante Múltipla e a Análise Logística.

5.1. Análise Discriminante Múltipla

Tal como referido ao longo deste trabalho, a Análise Discriminante Múltipla (ADM) é um método utilizado para designar técnicas estatísticas que permitem estudar as diferenças entre dois grupos já definidos, o grupo das empresas insolventes e o grupo das empresas saudáveis, com base num conjunto de características dos seus elementos, de modo a identificar e interpretar essas diferenças.

Visando o principal objetivo deste estudo, a elaboração de um modelo preditivo da situação de insolvência, a ADM corresponde à primeira análise que permitirá selecionar quais as variáveis que deverão ser incluídas nesse modelo final. Neste âmbito, esta análise encontra-se dividida em quatro etapas, nomeadamente na classificação das empresas como insolventes e saudáveis, na seleção das variáveis com base na análise dos coeficientes estandardizados, na análise da matriz de Correlação de *Pearson* e por fim, na Análise da Variância (ANOVA).

Deste modo, como primeira etapa, importa distinguir quais as empresas que são consideradas insolventes e saudáveis dentro da variável dicotómica *Insolvency*, de acordo com o conjunto de variáveis independentes definidas no Subcapítulo 4.2. As empresas insolventes assumem o valor de 1, enquanto as empresas saudáveis são classificadas com o valor de 0. Assim, resulta a *Confusion Table* que classifica cada empresa como insolvente ou saudável, com base nas variáveis explicativas em análise no presente estudo, avaliando a capacidade discriminante do modelo.

Tabela 4 – Classificação das Empresas – Confusion Table

Insolvency	Classificação		Total
	0	1	
0	6802 97,46%	177 2,54%	6979 100%
1	18 66,67%	9 33,33%	27 100%
Total	6820 97,39%	186 2,61%	7006 100%

Nota: Na Tabela 4 é apresentada a *Confusion Table*, que resulta do 1º passo da ADM que classifica as empresas que considera insolventes com valor de um (1) e as empresas consideradas como saudáveis com o valor de zero (0), de acordo com o conjunto de variáveis discriminantes descritas em detalhe no Subcapítulo 4.2.

Desta análise, é possível observar que 33,33% (97,46%) das empresas classificadas como insolventes (saudáveis) estão corretamente classificadas e que 66,67% (2,54%) das empresas classificadas como insolventes (saudáveis) não estão corretamente classificadas.

Da classificação proposta por Ohlson (1980), surge a Tabela 5, onde são destacados os Erros do Tipo I, isto é, a percentagem de empresas saudáveis classificadas como insolventes e os Erros do Tipo II, em que as empresas insolventes são classificadas como saudáveis.

Tabela 5 – Erros Tipo I e II derivados da Classificação das Empresas

Previsão do Resultado	Verdadeiro Positivo Empresas Insolventes (y=1)	Falso Positivo Erro tipo I: empresas saudáveis classificadas como insolventes
	Falso Negativo Erro tipo II: empresas insolventes classificadas como saudáveis	Verdadeiro Negativo Empresas saudáveis (y=0)

Nota: Na Tabela 5 são apresentados os Erros do Tipo I (falsos positivos ou empresas “saudáveis”, classificadas como “insolventes”) e Erros do Tipo II (falsos negativos ou empresas “insolventes” classificadas como “saudáveis”), de acordo com a interpretação de (Ohlson, 1980). Da análise da Tabela é possível perceber se as empresas estão corretamente classificadas, cruzando as empresas observadas com as empresas classificados pela aplicação da ADM.

As próximas etapas da ADM focam-se na seleção de variáveis que permitem reduzir a probabilidade de erros nos modelos de previsão de insolvência, visto que uma má classificação das empresas implica elevados custos, por exemplo, no caso do Erro do Tipo II em que as empresas que se encontram em estado de insolvência são classificadas como saudáveis. Para além disso, esta próxima análise contribui para melhorar a capacidade preditiva do modelo na identificação de potenciais empresas insolventes.

A Tabela 6 apresenta os coeficientes da função discriminante, não estandardizados e estandardizados. Os valores referentes aos coeficientes não estandardizados correspondem ao peso de cada variável explicativa, conferido pela capacidade preditiva na atribuição de cada observação ao grupo ao qual pertence - insolvente ou saudável (Sage, 2015). Contudo, a análise

dos coeficientes não estandardizados apresenta problemas de escala entre variáveis, por exemplo, no caso da variável Dimensão, medida como o logaritmo dos ativos, o seu indicador será superior ao dos rácios, conduzindo a um maior peso discriminativo desta variável. Para colmatar esta limitação, os coeficientes sofrem uma estandardização, permitindo identificar as variáveis que apresentam na realidade maior capacidade explicativa.

Os coeficientes estandardizados de cada variável devem ser interpretados mediante o seu valor absoluto, não considerando se o sinal do coeficiente é positivo ou negativo.

Tabela 6 – Análise dos Coeficientes da Função Discriminante

Variáveis	Coeficientes <i>Unstandardized</i>	Coeficientes <i>Standardized</i>
Fundo de Maneio	-1,0261	-0,2950
Rentabilidade Económica	0,8191	0,1558
Rentabilidade Vendas	0,8487	1,5272
ROA	1,8300	0,3314
Rentabilidade Ativo	-2,5678	-0,4928
VM Vendas	0,0808	1,3627
Margem EBIT	-0,4209	-0,6243
Liquidez Geral	-0,0085	-0,0161
Liquidez Imediata	-0,0363	-0,0398
Rotação Ativo	-0,0859	-0,0540
Inventário	0,0043	0,3493
Endividamento	-1,7120	-0,3794
Solvabilidade	-0,0004	-0,2475
Ganhos Retidos	-0,1166	-0,1157
Crescimento Vendas	-0,1958	-0,0839
Dimensão	-0,1331	-0,2493
Q(Tobin)	-0,0085	-0,0792
Constante	1,8300	-

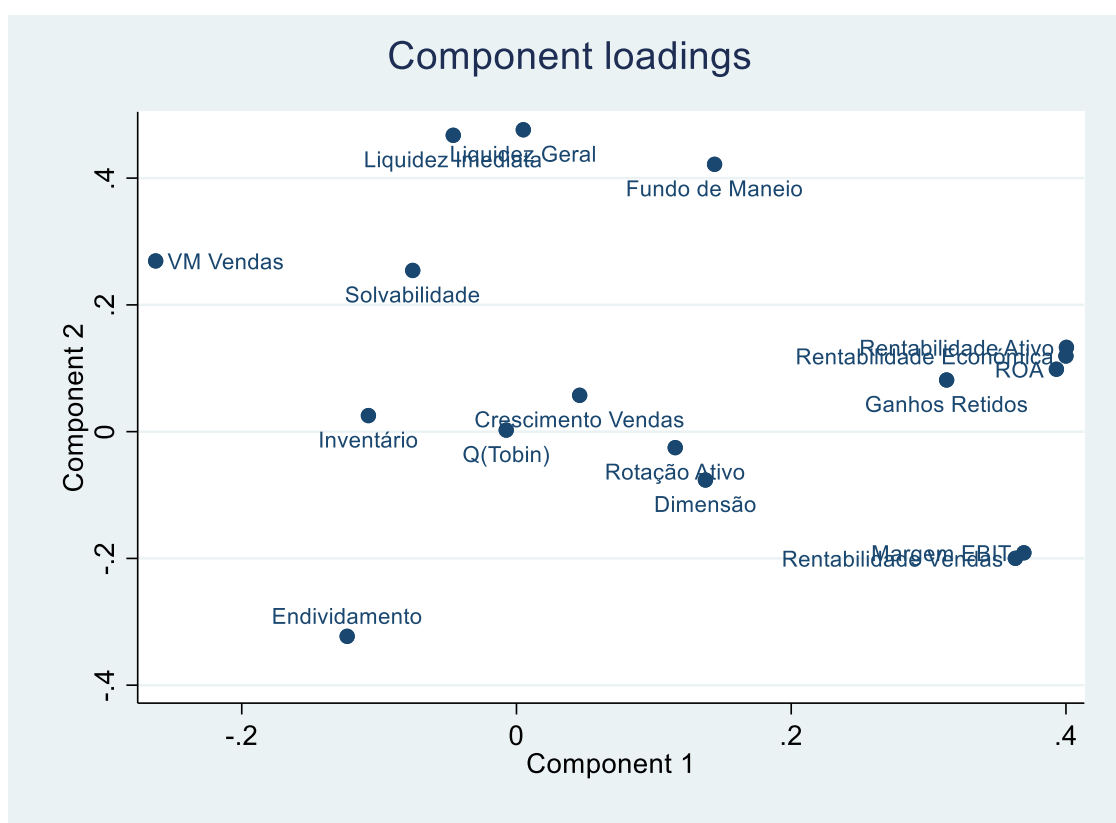
Nota: A Tabela 6 apresenta os coeficientes *unstandardized* e os coeficientes *standardized* que correspondem à 2ª etapa da ADM. As variáveis são as descritas detalhadamente no Subcapítulo 4.2.

Através da análise dos coeficientes estandardizados, são de salientar as variáveis Fundo de Maneio, Rentabilidade Vendas, ROA, Rentabilidade Ativo, VM Vendas, Margem EBIT, Inventário, Endividamento e Dimensão, que são as que apresentam coeficientes absolutos de maior magnitude e, conseqüentemente, demonstram maior peso na discriminação entre empresas insolventes e saudáveis. Destaca-se ainda o facto de que, por norma, o *cutoff point* considerado razoável como determinante na identificação de variáveis com maior ou menor poder

discriminatório é de 30% (Sage, 2015). No entanto, é de reforçar que as variáveis Solvabilidade e Dimensão encontram-se bastante perto da zona de *cutoff*.

De forma a sustentar os resultados obtidos com a tabela anterior, é possível observar no Gráfico 1 a representação dos coeficientes e, assim, verificar a possível sobreposição destes, revelando um grau superior de correlação entre as variáveis, maior do que o desejado. Neste caso, é observável que as variáveis Margem EBIT e Rentabilidade Vendas se encontram sobrepostas, assim como a Rentabilidade Económica, a Rentabilidade do Ativo e a ROA, o que evidencia um elevado grau de correlação entre as mesmas, dificultando a diferenciação entre grupos. Isto conduz à opção de estas variáveis não serem incluídas em simultâneo no modelo final de previsão.

Gráfico 1 – Coeficientes *Standardized*



Nota: O Gráfico 1 apresenta o peso de cada variável em função dos coeficientes *standardized*. As variáveis são as descritas detalhadamente no Subcapítulo 4.2.

Esta análise constitui apenas uma das etapas da ADM, sendo que estes resultados poderão sofrer alterações mediante as análises que se seguem. Neste sentido, a próxima etapa é a análise de correlação de *Pearson*, *Structure Matrix*, que avalia a correlação entre grupos em função das variáveis explicativas.

A partir da análise da Tabela 7, é possível determinar o grau de correlação entre o peso de cada variável e a função discriminante, através dos coeficientes. Tal como na etapa anterior, estes valores devem ser interpretados em função do seu valor absoluto, ou seja, as variáveis com maior magnitude revelam-se melhores preditores, considerando-se, de igual forma, um *cutoff point* de 30% (Sage, 2015).

Tabela 7 - Análise de Correlação de Pearson (*Structure Matrix*)

Variáveis	Coefficientes
Fundo de Maneio	-0,0903
Rentabilidade Económica	-0,1192
Rentabilidade Vendas	-0,2422
ROA	-0,1002
Rentabilidade Ativo	-0,1220
VM Vendas	0,6656
Margem EBIT	-0,2306
Liquidez Geral	0,0003
Liquidez Imediata	0,0100
Rotação Ativo	-0,1286
Inventário	0,2629
Endividamento	-0,1775
Solvabilidade	-0,0442
Ganhos Retidos	-0,1970
Crescimento Vendas	-0,0630
Dimensão	-0,3105
Q(Tobin)	-0,0221

Nota: A Tabela 7, denominada por *Structure Matrix*, reporta os coeficientes que medem a correlação entre cada variável explicativa e a função discriminante. As variáveis são as descritas detalhadamente no Subcapítulo 4.2.

De acordo com os critérios definidos e avaliando os dados da tabela, as variáveis explicativas que se destacam são: Rentabilidade Vendas, VM Vendas, Margem EBIT, Inventário, Endividamento, Ganhos Retidos e Dimensão. Em suma, os resultados obtidos com esta análise reforçam a anterior, apesar de se observarem ligeiras alterações.

De seguida, recorre-se à Análise de Variância (ANOVA), que permite avaliar perceções sobre as médias das populações. Esta técnica canónica avalia a seleção das variáveis mediante o teste estatístico F sob a hipótese nula (H_0) de inexistência de diferenças significativas relativamente ao poder explicativo das variáveis, sendo esta rejeitada para um grau de significância ao nível de 5%.

Tabela 8 – Análise de Variância (ANOVA)

Variáveis	F-statistic test	P-value	
Fundo de Maneio	1,9066	0,1674	
Rentabilidade Económica	3,3221	0,0684	*
Rentabilidade Vendas	13,7210	0,0002	***
ROA	2,3482	0,1255	
Rentabilidade Ativo	3,4796	0,0622	*
VM Vendas	103,6500	0,0000	***
Margem EBIT	12,4380	0,0004	***
Liquidez Geral	1,7e-05	0,9967	
Liquidez Imediata	0,0232	0,8790	
Rotação Ativo	3,8677	0,0493	**
Inventário	16,1730	0,0001	***
Endividamento	7,3752	0,0066	***
Solvabilidade	0,4566	0,4993	
Ganhos Retidos	9,0766	0,0026	***
Crescimento Vendas	0,9284	0,3353	
Dimensão	22,5590	0,0000	***
Q(Tobin)	0,1145	0,7350	

Nota: A Tabela 8 apresenta os resultados da análise de variância ANOVA para cada variável explicativa. As variáveis são as descritas detalhadamente no Subcapítulo 4.2. Os coeficientes são considerados estatisticamente significativos ao nível de 1% (***), 5% (**) e 10% (*).

Com base no *output* gerado na Tabela 8, as variáveis significativas ao nível de, pelo menos, 5%, isto é, que apresentam elevada capacidade explicativa da situação de insolvência são: Rentabilidade Vendas, VM Vendas, Margem EBIT, Rotação Ativo, Inventário, Endividamento, Ganhos Retidos e Dimensão.

Em suma, considerando os resultados obtidos nas diferentes etapas da ADM, as variáveis que mais se destacaram devido ao seu poder explicativo foram: Rentabilidade Vendas, VM Vendas, Margem EBIT, Inventário, Endividamento, Ganhos Retidos e Dimensão.

5.2. Análise Logística

Apesar das vantagens associadas à ADM, a sua validade e robustez foi alvo de críticas por diversos autores, entre os quais Ohlson (1980), que propôs um modelo logístico para avaliar a previsão da insolvência financeira. A Análise *Logit*, contrariamente à ADM, não exige que as variáveis explicativas apresentem uma distribuição normal, nem que os grupos sejam discretos e identificáveis.

Neste âmbito, e considerando o objetivo deste trabalho de formular um modelo semelhante ao Modelo *Logit* (Ohlson, 1980) que revele uma capacidade de prever a situação de insolvência até cinco anos antes do evento, é apresentado o seguinte modelo definido na equação 4:

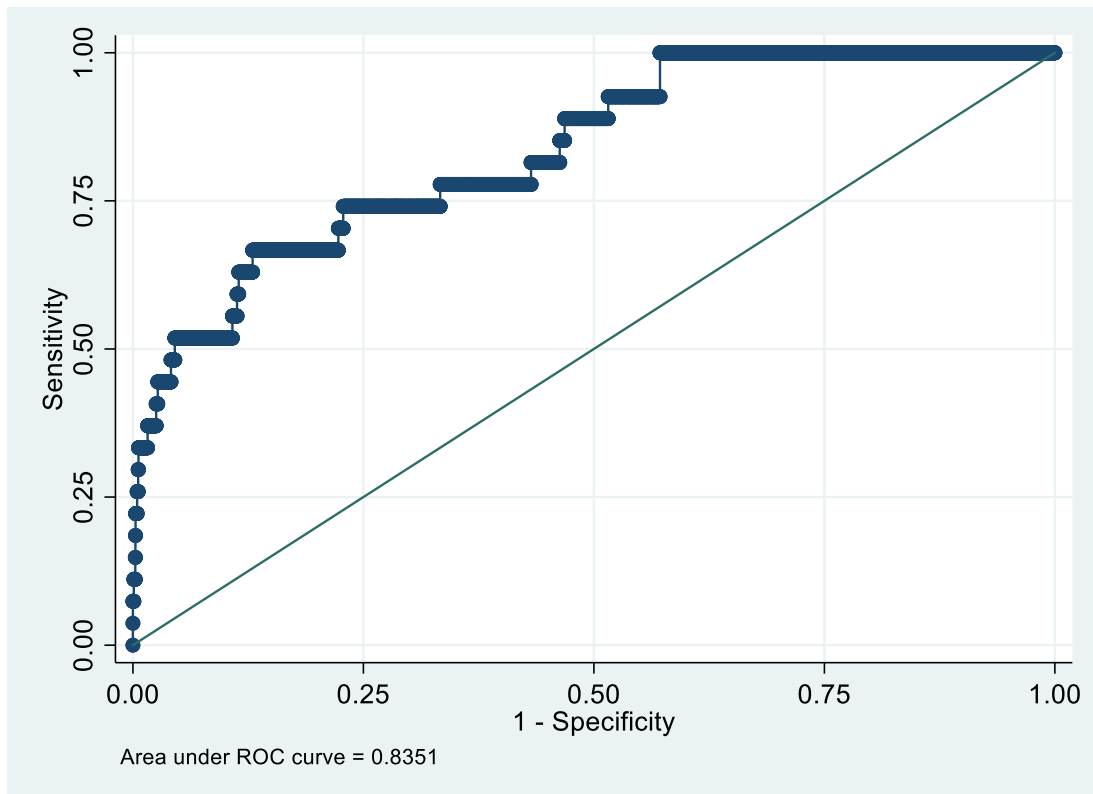
$$Prob(Insolvency = 1) = \beta'X_{i,t=-1} + \lambda_j + \gamma_t + \varepsilon_{it} \quad (4)$$

A variável dependente é a *Insolvency*, de natureza binária, assumindo o valor 1 quando a empresa é insolvente e 0 caso contrário. O vetor $X_{i,t=-1}$ representa as variáveis independentes desfasadas um ano. Quanto aos parâmetros λ_j e γ_t , estes controlam a heterogeneidade não observada, relativamente à variação temporal e à indústria, e o termo ε_{it} representa o erro da regressão.

Dentro desta análise, o primeiro passo consiste na classificação das empresas mediante o seu *status*, insolvente ou saudável. Para tal, recorre-se a um gráfico, *Receiver Operating Characteristic Curve* (ROC), que permite avaliar o desempenho de classificação binária à medida que o seu limiar de discriminação varia consoante as variáveis independentes do modelo (*covariates*).

A Curva ROC apresenta a taxa de casos verdadeiros positivos (*Sensitivity*) versus a taxa de casos falsos positivos, para diferentes pontos de *cutoff*. Para além destas, a taxa de casos de verdadeiros negativos versus a taxa de falsos negativos designa-se por *Specificity*. Neste sentido, pretende-se identificar o ponto de *cutoff* que minimiza o risco de se incorrer em erros: Erros do Tipo I que correspondem à taxa de falsos positivos, expressa por (1- *Specificity*) e Erros do Tipo II medidos pela taxa de falsos negativos, expressa por (1-*Sensitivity*).

Gráfico 2 – ROC (Receiver Operating Characteristic Curve)



Nota: O Gráfico 2 apresenta a representação gráfica da Curva ROC. A “Area under ROC curve” representa uma medida de discriminação que avalia entre 0 e 1 a precisão com que as empresas insolventes estão corretamente classificadas, de acordo com as variáveis descritas detalhadamente no Subcapítulo 4.2.

A partir da análise do Gráfico 2, é possível observar qual a previsão com que a variável *Insolvency* identifica de forma correta as empresas insolventes através da “Area under ROC curve”. Neste sentido, obteve-se um nível de precisão de cerca de 83,51% o que, de modo geral, representa um resultado satisfatório, considerando que a capacidade preditiva de um modelo deve superar os 70%, de acordo com a maioria dos estudos anteriores (Hosmer Jr, Lemeshow, & Sturdivant, 2013).

De seguida, é essencial recorrer à estandardização dos coeficientes da estimação do Modelo *Logit* para melhorar a análise comparativa entre a magnitude dos coeficientes e o seu poder explicativo.

Na Tabela 9 constam os resultados da estimação da equação (4) antes da aplicação da estandardização dos coeficientes.

Tabela 9 - Resultados Antes do Processo de Estandarização das Variáveis

Variável dependente: <i>Insolvency</i>	Coefficientes	D.P.	Z-statistic test	P-value	
Fundo de Maneio	-0,8932	0,8242	-1,0800	0,2780	
Rentabilidade Económica	-0,2998	3,5112	-0,0900	0,9320	
Rentabilidade Vendas	-0,0191	0,4481	-0,0400	0,9660	
ROA	2,5743	2,4946	1,0300	0,3020	
Rentabilidade Ativo	-2,1074	2,5761	-0,8200	0,4130	
VM Vendas	0,0234	0,0050	4,6600	0,0000	***
Margem EBIT	0,2101	0,5554	0,3800	0,7050	
Liquidez Geral	-0,0351	0,1644	-0,2100	0,8310	
Liquidez Imediata	0,1048	0,2796	0,3700	0,7080	
Rotação Ativo	-0,2216	0,3648	-0,6100	0,5430	
Inventário	0,0032	0,0015	2,1100	0,0340	**
Endividamento	-4,8198	1,5715	-3,0700	0,0020	***
Solvabilidade	-0,0063	0,0061	-1,0300	0,3040	
Ganhos Retidos	-0,0888	0,1506	-0,5900	0,5550	
Crescimento Vendas	0,0568	0,3656	0,1600	0,8760	
Dimensão	-0,4253	0,0548	-7,7700	0,0000	***
Q(Tobin)	-0,0308	0,0206	-1,4900	0,1350	

Observações	7006
Significância Global	0,0000
Coeficientes: Chi2-test	

Nota: Na Tabela 9 reportam-se os resultados da estimação da regressão (4) antes do processo de estandarização dos coeficientes. A variável dependente é uma *dummy* que classifica as empresas com um (1) se insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. É, ainda, apresentado o N° de Observações e a Significância Global dos Coeficientes. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (***) e 5% (**).

Os coeficientes reportados na tabela acima são apresentados como *log-odds*, emergindo a necessidade de serem convertidos para probabilidades marginais de forma a proceder-se à sua interpretação, de acordo com a seguinte expressão:

$$P = \frac{e^{\log-odds}}{1+e^{\log-odds}} \quad (5)$$

Através desta equação, mediante o sinal do coeficiente das variáveis que constituem o modelo, é possível aferir quais as variáveis com maior ou menor probabilidade de insolvência. Neste sentido, e a título exemplificativo, quanto maior for o valor da variável Inventário, maior será a probabilidade de insolvência. Por outro lado, quando o valor da variável Liquidez Geral sofre acréscimos, a probabilidade de insolvência é menor.

No caso da variável Liquidez Geral, o aumento de uma unidade da mesma resultará na diminuição de aproximadamente 0,0351 na probabilidade logarítmica da variável dependente, *Insolvency*, mantendo as restantes variáveis explicativas constantes⁵.

De acordo com os resultados obtidos das análises realizadas até ao momento, foram selecionadas as variáveis que demonstraram uma maior capacidade explicativa da situação de insolvência. Posteriormente, foi realizada a estandardização dos coeficientes, cujos resultados são reportados na Tabela 10. O objetivo primordial desta técnica é mitigar a diferença de escala verificada entre as variáveis, conduzindo a resultados mais robustos. É observável uma diferença significativa ao nível da magnitude dos coeficientes, sendo que, após o processo de estandardização, os mesmos variam entre 0 e 1, com exceção da variável Margem EBIT, que apresenta um coeficiente superior a 1. Tal resultado deriva do seu elevado grau de correlação com a variável Rentabilidade Vendas, o que foi previamente detetado na ADM. Assim, esta variável deverá ser excluída do modelo final.

Tabela 10 - Resultados Após o Processo de Estandarização das Variáveis

Variável dependente: <i>Insolvency</i>	Coefficientes	D.P.	Z-statistic test	P-value	
Fundo de Maneio	0,1959	0,0333	5,8800	0,0000	***
Rentabilidade Vendas	-1,9125	0,3225	-5,9300	0,0000	***
ROA	0,2063	0,0824	2,5000	0,0120	**
Rentabilidade Ativo	-0,3713	0,0918	-4,0400	0,0000	***
VM Vendas	0,2964	0,0696	3,7900	0,0000	***
Margem EBIT	2,1150	0,3043	6,9500	0,0000	***
Rotação Ativo	-0,1628	0,0270	-6,0400	0,0000	***
Inventário	-0,0558	0,0270	-2,0700	0,0390	**
Endividamento	-0,3561	0,0315	-11,3200	0,0000	***
Solvabilidade	-0,1312	0,0277	-4,7300	0,0000	***
Ganhos Retidos	-0,4394	0,0502	-8,7500	0,0000	***
Dimensão	-0,2056	0,0266	-7,7400	0,0000	***
Observações		8234			
Significância Global		0,0000			
Coeficientes: Chi2-test					

Nota: Na Tabela 10 apresentam-se os resultados da estimação da regressão (4) após o processo de estandardização dos coeficientes. A variável dependente é uma *dummy* que classifica as empresas com um (1) se insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4. 2. É, ainda, apresentado o N° de Observações e a Significância Global dos Coeficientes. Os coeficientes são estatisticamente significativos ao nível de pelo menos 1% (***) e 5% (**).

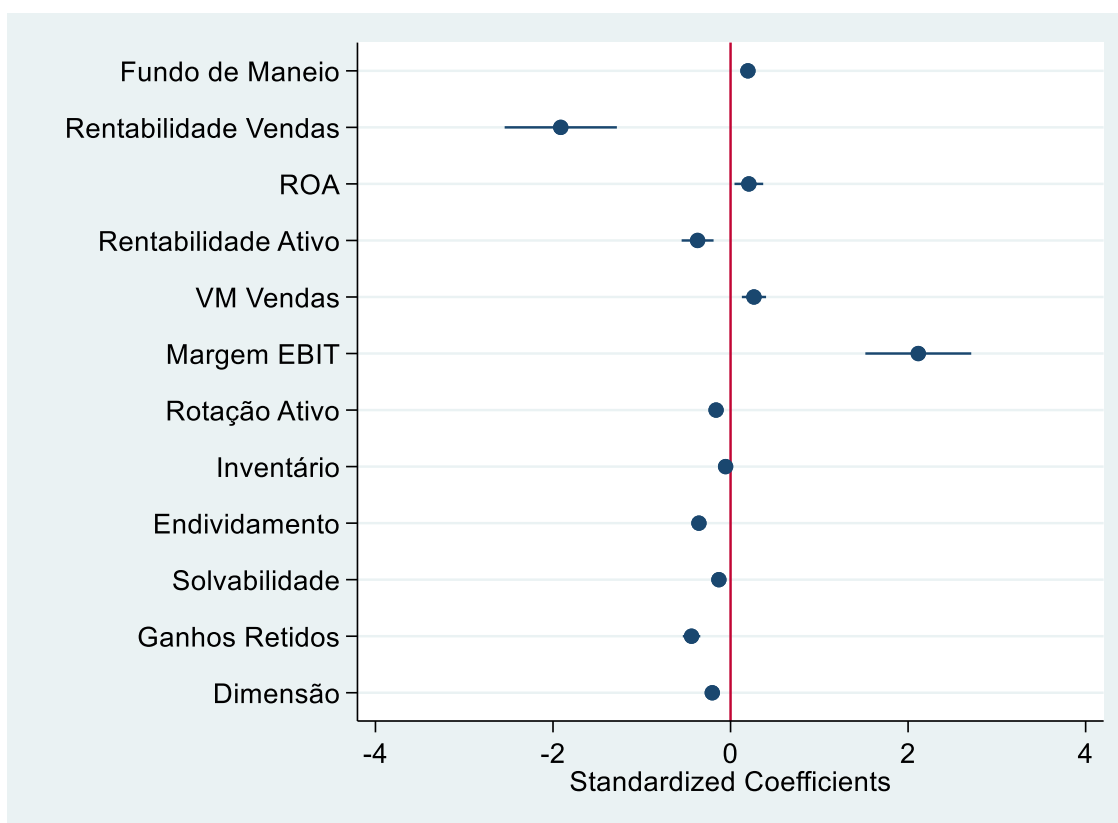
Contrariamente à Tabela 9, em que apenas as variáveis VM Vendas, Endividamento e Dimensão apresentavam significância estatística para o nível de 1%, na Tabela 10 todas as

⁵ A probabilidade de insolvência é obtida através da equação (5): $P = \frac{e^{0,0351}}{1+e^{0,0351}} = 0,5088$, implicando que a variação de 1% no rácio da Liquidez Geral corresponde, de grosso modo, à variação de 50,88% na probabilidade de insolvência, mantendo as restantes variáveis constantes.

variáveis são estatisticamente significativas ao nível de 1%, excetuando as variáveis ROA e Inventário, que são significativas ao nível de 5%.

A análise dos coeficientes estandardizados é também suportada pela representação gráfica da significância dos mesmos. O Gráfico 3 possibilita identificar quais os coeficientes que se encontram mais afastados de zero (0), o que significa que são estatisticamente diferentes de zero.

Gráfico 3 – Significância dos Coeficientes Estandardizados



Nota: O Gráfico 3 representa a significância estatística das variáveis independentes em análise e descritas detalhadamente no Subcapítulo 4.2, após o processo de estandardização dos coeficientes.

De acordo com o Gráfico 3, a variável que se encontra mais próxima de 0 é a Inventário, sendo esta uma variável que, após o processo de estandardização, não apresenta significância estatística ao nível de 1%.

Após esta análise, todas as variáveis que sofreram estandardização revelam um elevado poder explicativo da situação de insolvência, sendo estas: Fundo de Maneio, Rentabilidade Vendas, ROA, Rentabilidade Ativo, VM Vendas, Margem EBIT, Rotação Ativo, Inventário, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão. Comparando estes resultados com os obtidos na ADM, é possível identificar as variáveis com maior capacidade preditiva, visto que as que se destacaram

devido ao seu elevado poder estatístico na ADM tiveram o mesmo comportamento quando sujeitas à Análise Logística.

No entanto, o resultado obtido com a Análise Logística comporta maior relevância e deve sobrepor-se aos da ADM, no caso de se verificarem ligeiras discrepâncias. Para além disso, a variável Margem EBIT é uma variável com elevado grau de correlação, ou seja, apesar de ser estatisticamente significativa em ambas as análises, a mesma não será integrada no modelo final, assim como as variáveis Inventário e ROA que, tal como previamente mencionado, após a estandardização dos coeficientes são as únicas que não apresentam significância estatística para um nível de 1%.

Posto isto, as variáveis selecionadas para integrarem o modelo final são: Rentabilidade Vendas, VM Vendas, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão.

6. Apresentação e Discussão dos Resultados

Neste capítulo será apresentado o modelo final de previsão de insolvência baseado no Modelo *Logit* proposto por Ohlson (1980), sendo que as variáveis incluídas foram previamente selecionadas com recurso a duas análises, a ADM e a Análise Logística, de forma a averiguar a sua capacidade preditiva da situação de insolvência.

6.1. Modelo *Logit* com capacidade preditiva entre um e cinco anos antes do evento de insolvência

De acordo com os resultados obtidos no capítulo anterior, as variáveis que integrarão o modelo são: Rentabilidade Vendas, VM Vendas, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão. A equação 6 apresenta a estimação desse modelo, em que a variável dependente, *Insolvency*, é de natureza binária, assumindo o valor de 1 quando as empresas estão classificadas como insolventes (*Insolvency*=1) e 0, caso contrário.

$$\begin{aligned} Prob(Insolvency = 1) = & \\ & \alpha + \beta_1 \text{Rentabilidade Vendas} + \beta_2 \text{VM Vendas} \\ & + \beta_3 \text{Endividamento} + \beta_4 \text{Solvabilidade} + \beta_5 \text{Ganhos Retidos} \\ & + \beta_6 \text{Dimensão} + \varepsilon_{it} \end{aligned} \quad (6)$$

A Tabela 11 apresenta os resultados da estimação dos coeficientes da regressão logística definida pela equação 6, com defasamentos de 1, 2, 3, 4 e 5 anos antes da insolvência. A partir do teste de significância conjunta, qui-quadrado, é possível aferir que as variáveis explicativas que incorporam o modelo apresentam elevada capacidade preditiva da situação de insolvência para um nível de significância de 1% até cinco anos antes. Tal significa que o conjunto de variáveis independentes sob análise apresenta significância estatística conjunta e é capaz de explicar o comportamento da variável dependente.

Tabela 11 – Resultados da Estimação do Modelo *Logit* 1, 2, 3, 4 e 5 Anos Antes da Insolvência

Variável dependente: <i>Insolvency</i> Modelo	1 Ano (1)		2 Anos (2)		3 Anos (3)		4 Anos (4)		5 Anos (5)	
	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Rentabilidade Vendas	0,1026	0,0320	0,1405	0,0020	0,1184	0,0110	0,1525	0,0060	0,0857	0,0000
VM Vendas	0,0241	0,0000	0,0241	0,0000	0,0222	0,0000	0,0226	0,0000	0,0190	0,0000
Endividamento	-4,4458	0,0000	-4,4046	0,0010	-4,2530	0,0030	-3,5664	0,0050	-2,8424	0,0070
Solvabilidade	-0,0079	0,1090	-0,0060	0,0640	-0,0054	0,0600	-0,0109	0,0810	-0,0334	0,0620
Ganhos Retidos	-0,1674	0,0480	-0,1883	0,0270	-0,1903	0,0340	-0,1420	0,1520	-0,0470	0,6480
Dimensão	-0,1469	0,3490	-0,1302	0,4070	-0,1007	0,5280	-0,2331	0,1600	-0,4296	0,0010
Constante	-3,2445	0,0490	-3,5101	0,0310	-3,9155	0,0180	-2,6833	0,1170	-0,7434	0,5820
Observações	8500		8498		8496		8492		8488	
Significância Global Coeficientes: Chi2-test	0,0000		0,0000		0,0000		0,0000		0,0000	
Pseudo R ²	0,1189		0,1067		0,0825		0,0828		0,0781	
Cutoff point	0,0029		0,0027		0,0029		0,0023		0,0020	
Obs. corretamente classificadas	63,48%		60,87%		65,36%		67,02%		71,69%	
Curva ROC	75,99%		74,49%		73,05%		75,30%		81,43%	

Nota: A Tabela 11 apresenta os resultados da estimação da equação (6), nomeadamente os coeficientes e o respetivo *p-value*. A variável dependente é uma *dummy* que classifica as empresas com um (1) se insolvente e zero (0), caso contrário. As variáveis independentes são as descritas detalhadamente no Subcapítulo 4.2. A variável dependente é binária, assumindo o valor de um (1) quando a empresa é insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. É, ainda, apresentado o N° de Observações, a Significância Global dos Coeficientes, o Pseudo R², o *Cutoff point* e a percentagem de Observações Corretamente Classificadas, para 1, 2, 3, 4 e 5 anos antes da insolvência. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% encontram-se assinalados a negrito. Os erros-padrão são *White-robust*.

Da análise dos cinco modelos da tabela acima, conclui-se que apenas os coeficientes das variáveis Ganhos Retidos e Dimensão apresentam o sinal esperado, sinal negativo, ou seja, estão negativamente correlacionadas com a probabilidade de insolvência. No entanto, as variáveis Rentabilidade Vendas, VM Vendas, Endividamento e Solvabilidade apresentam sinais contrários ao esperado. No entanto, é mais importante o contributo conjunto de todas as variáveis para a estimativa de probabilidade de insolvência, do que a análise isolada de cada variável (Ohlson, 1980).

Relativamente à significância estatística das variáveis, a mesma não é constante para os cinco modelos. As variáveis Rentabilidade Vendas, VM Vendas e Endividamento são estatisticamente significativas para um nível de significância de, pelo menos, 5% para todos os anos em análise, antes da ocorrência da insolvência. A variável Solvabilidade apresenta significância estatística para um nível de 10% nos modelos (2), (3), (4) e (5). Já a variável Ganhos Retidos é estatisticamente significativa ao nível 5% nos modelos (1), (2) e (3). Por fim, a estimativa do coeficiente da variável Dimensão só apresenta significância estatística para um nível de 1%, no modelo (5).

Além disso, é importante salientar o contributo global do modelo que permitirá auferir acerca da probabilidade de uma empresa entrar em situação de insolvência, sendo que, em média, 63,48% e 71,69% das observações estão corretamente classificadas para um e cinco anos antes da insolvência, respetivamente. Adicionalmente, no que concerne à capacidade de discriminação entre empresas insolventes e saudáveis, o modelo apresenta um nível de precisão médio de, aproximadamente, 76% (Curva ROC), para os cinco anos.

Posteriormente, a equação 6 foi reestimada de modo a averiguar o impacto que, além do horizonte temporal, a inclusão de um novo parâmetro tem sobre a insolvência de uma empresa. Neste sentido, foram incluídos parâmetros para controlar os efeitos temporais e de indústria.

A Tabela 12 dispõe então os resultados da estimação da regressão logística definida na equação 6, por indústria e com efeitos temporais.

Tabela 12 – Resultados da Estimação do Modelo Logit com Controlo para Efeitos Temporais e Indústria

Variável dependente: <i>Insolvency</i> Modelo	1 Ano		2 Anos		3 Anos		4 Anos		5 Anos	
	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value	Coef.	P-value
Rentabilidade Vendas	0,1159	0,1640	0,2199	0,0080	0,1962	0,0250	0,3303	0,0020	0,2528	0,1390
VM Vendas	0,0405	0,0000	0,0466	0,0000	0,0457	0,0000	0,0607	0,0000	0,0534	0,0060
Endividamento	-5,3501	0,0000	-5,5397	0,0010	-5,1582	0,0010	-4,1554	0,0020	-2,9305	0,0060
Solvabilidade	-0,0060	0,0030	-0,0053	0,0010	-0,0051	0,0010	-0,0079	0,0010	-0,0181	0,0340
Ganhos Retidos	-0,1817	0,0960	-0,2009	0,0560	-0,2085	0,0520	-0,1305	0,2510	-0,0322	0,7780
Dimensão	-0,1445	0,3860	-0,1075	0,5170	-0,0733	0,6590	-0,1886	0,2660	-0,3539	0,0140
Constante	-5,9620	0,0020	-6,4930	0,0010	-6,8410	0,0000	-6,3735	0,0000	-4,2582	0,0060
Observações	4986		4696		4333		4207		3848	
Significância Global	0,0000		0,0000		0,0000		0,0000		0,0000	
Coeficientes: Chi2-test	0,3044		0,3022		0,2823		0,2955		0,2647	
Pseudo R ²	0,3044		0,3022		0,2823		0,2955		0,2647	
Cutoff point	0,0088		0,0096		0,0087		0,0099		0,0097	
Obs. corretamente classificadas	85,40%		85,86%		83,54%		86,62%		88,46%	
Curva ROC	94,30%		93,65%		92,94%		93,73%		93,52%	

Nota: A Tabela 12 reporta os resultados da estimação da equação (6) - incluindo os parâmetros para controlo de efeitos temporais e indústria - nomeadamente, os coeficientes e o *p-value* associado a cada coeficiente. A variável dependente é binária, assumindo o valor de um (1) quando a empresa é insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. É, ainda, apresentado o N° de Observações, a Significância Global dos Coeficientes, o Pseudo R², o *Cutoff point* e a percentagem de Observações Corretamente Classificadas, para 1, 2, 3, 4 e 5 anos antes da insolvência. Os coeficientes estatisticamente significativos ao nível de pelo menos 10% encontram-se assinalados a negrito. Os erros-padrão são *White-robust*.

Na Tabela 12 é possível observar que o sinal e a magnitude dos coeficientes não sofreu alteração em relação aos valores obtidos na Tabela 11, embora a significância estatística de alguns coeficientes tenha sido afetada, nomeadamente a variável Rentabilidade Vendas que deixou de ser

significativa nos modelos (1) e (5). Contrariamente, a variável Solvabilidade passou a apresentar significância estatística em todos os modelos, com um nível de 1% para os modelos (1), (2), (3) e (4) e, de 5% no modelo (5).

Relativamente à significância global, o modelo obtido a partir da equação 6 é significativamente estatístico para o nível de significância de 1% até cinco anos antes do evento de insolvência, o que traduz a boa capacidade preditiva das variáveis incluídas no modelo na explicação da situação de insolvência.

De destacar o aumento das observações que estão corretamente classificadas em relação aos resultados apresentados na Tabela 11, que varia entre 85,40% e 88,46% para um e cinco anos antes da insolvência, respetivamente. Adicionalmente, no que concerne à capacidade de discriminação entre empresas insolventes e saudáveis, o modelo apresenta um nível de precisão médio superior ao da Tabela 11, aproximadamente 94% (Curva ROC) para os cinco anos. Destaca-se ainda que, com a introdução destes parâmetros, verificou-se uma melhoria do nível da qualidade de ajustamento do modelo, observável através do aumento do valor do Pseudo R^2 da Tabela 11 para a Tabela 12.

6.2. Modelo Final de Previsão de Insolvência

Na sequência das análises anteriores, o modelo final de previsão de insolvência que engloba as variáveis previamente selecionadas é o definido pela equação 6. A Tabela 13 apresenta os resultados da estimação do modelo.

¶ Pseudo R^2 é um indicador estatístico que permite aferir acerca da qualidade de ajustamento do modelo que varia entre 0 e 1. No máximo assume o valor de 1, em caso de total adequação do modelo e no mínimo 0, caso contrário (Ohlson, 1980).

Tabela 13 - Resultados da Estimação do Modelo *Logit* Final

Variável dependente: <i>Insolvency</i>	Coefficientes	D.P.	Z-statistic test	P-value	
Rentabilidade Vendas	0,1281	0,0490	2,6100	0,0090	***
VM Vendas	0,0243	0,0031	7,8300	0,0000	***
Endividamento	-4,9812	1,2760	-3,9000	0,0000	***
Solvabilidade	-0,0082	0,0044	-1,8400	0,0650	*
Ganhos Retidos	-0,1651	0,0820	-2,0100	0,0440	**
Dimensão	-0,3127	0,1687	-1,8500	0,0640	*
Constante	-1,2467	1,7296	-0,7200	0,4710	

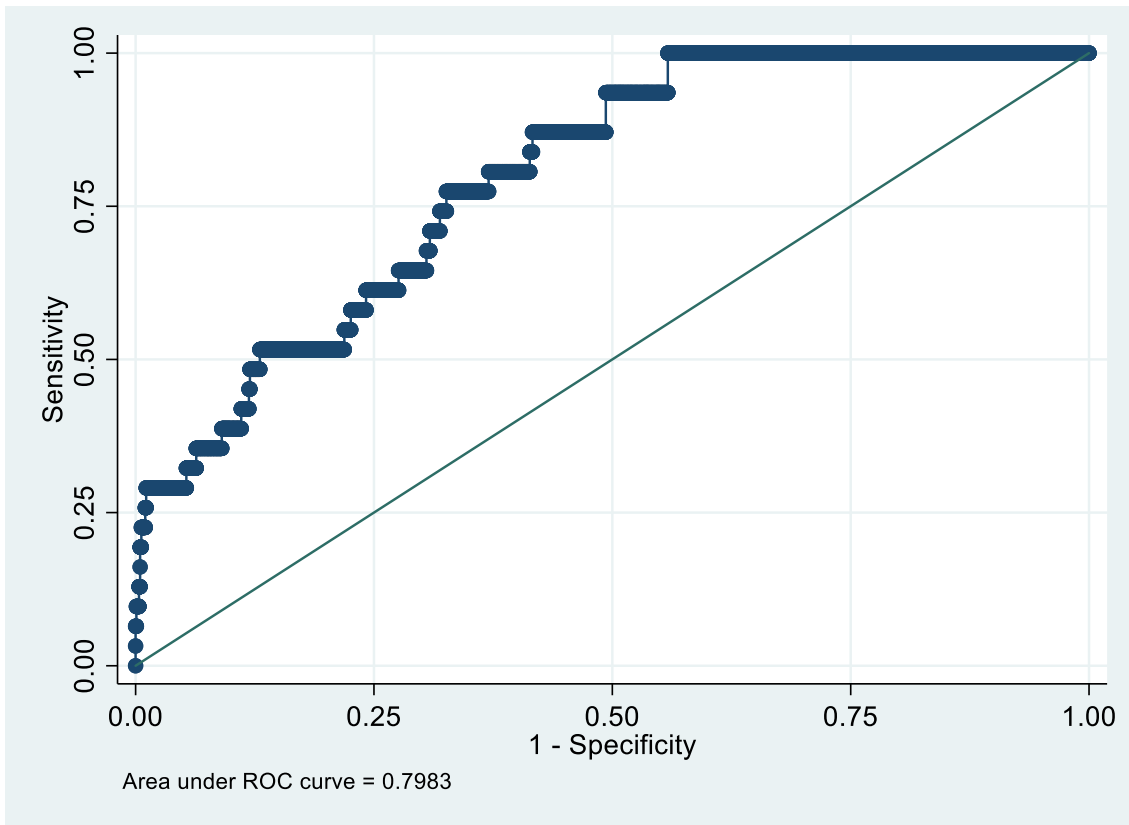
Observações	8504
Significância Global	0,0000
Coefficientes: Chi2-test	
Pseudo R ²	0,1382
Cutoff point	0,0033
Obs. corretamente classificadas	67,43%

Nota: A Tabela 13 apresenta os coeficientes da estimação da equação (6). A variável dependente é uma *dummy*, assumindo o valor um (1) se insolvente e zero (0) caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. A variável dependente é binária, assumindo o valor de um (1) quando a empresa é insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. É, ainda, apresentado o N° de Observações, a Significância Global dos Coeficientes, o Pseudo R², o *Cutoff point* e a percentagem de Observações Corretamente Classificadas. Os coeficientes são estatisticamente significativos ao nível de 1% (***) , 5% (**) e 10% (*). Os erros-padrão são *White-robust*.

Da análise da Tabela 13 conclui-se que a magnitude dos coeficientes é similar aos resultados obtidos com as análises anteriores e que todos os coeficientes são estatisticamente significativos.

O modelo geral apresenta uma percentagem de observações corretamente classificadas de 67,43%, sendo este valor próximo do comumente aceite (Hosmer Jr, Lemeshow, & Sturdivant, 2013) de, aproximadamente, 70%. Adicionalmente, tal como se pode observar no Gráfico 4 que representa a Curva ROC, o nível de precisão do modelo é de aproximadamente 79,83%. Tais resultados permitem concluir que, em média, o nível de precisão deste modelo revela um resultado satisfatório.

Gráfico 4 - Curva ROC do Modelo *Logit* Final



Nota: O Gráfico 4 apresenta a representação gráfica da Curva ROC. A “*Area under ROC curve*” representa uma medida de discriminação que avalia entre 0 e 1 a precisão com que as empresas insolventes estão corretamente classificadas.

Assim, a equação relativa ao modelo final de previsão de insolvência é a seguinte:

$$\begin{aligned}
 \text{Score} = & -1,2467 + 0,1281 \text{ Rentabilidade Vendas} \\
 & + 0,0243 \text{ VM Vendas} - 4,9812 \text{ Endividamento} \\
 & - 0,0082 \text{ Solvabilidade} - 0,1651 \text{ Ganhos Retidos} \\
 & - 0,3127 \text{ Dimensão}
 \end{aligned} \tag{7}$$

A equação (7) corresponde a um *Score* que deve ser convertido numa probabilidade. Deste modo, a probabilidade de a empresa entrar em insolvência é obtida através da seguinte expressão:

$$\text{Prob}(\text{Insolvency} = 1) = \frac{1}{1 + e^{-\text{Score}}} \tag{8}$$

Uma empresa é classificada como insolvente quando o valor do *Score* é superior ao valor do *cutoff point*, ou seja, uma vez que o ponto de *cutoff* é de 0,4%, uma empresa é insolvente se o *score* for superior a 0,4%.

6.3. Teste de Robustez

Com o propósito de testar a robustez dos resultados obtidos, desenvolveu-se um teste que pretende mimetizar um teste *Out-of-Sample*. Desta forma, pretende-se apurar a capacidade preditiva do modelo a partir da estimação dos coeficientes de uma subamostra que engloba o período de 2006 até 2013, inclusive, aplicados na outra metade da amostra, entre 2014-2021. A Tabela 14 apresenta os resultados desta análise.

Tabela 14 – Resultados da Estimação do Teste *Out-of-Sample*

Variável dependente: <i>Insolvency</i>	Coefficientes	D. P.	Z-statistic test	P-value
Rentabilidade Vendas	0,2661	0,3226	0,8300	0,4090
VM Vendas	-0,1168	0,1234	-0,9500	0,3440
Endividamento	-3,4766	1,3657	-2,5500	0,0110 **
Solvabilidade	-0,0639	0,0474	-1,3500	0,1770
Ganhos Retidos	0,0546	0,1091	0,5000	0,6170
Dimensão	-0,5288	0,1425	-3,7100	0,0000 ***
Constante	1,1481	1,4952	0,7700	0,4430

Observações	6340
Significância Global Coeficientes:	
Chi2-test	0,0000
Pseudo R^2	0,0920
Cutoff point	0,0031
Obs. corretamente classificadas	72,00%

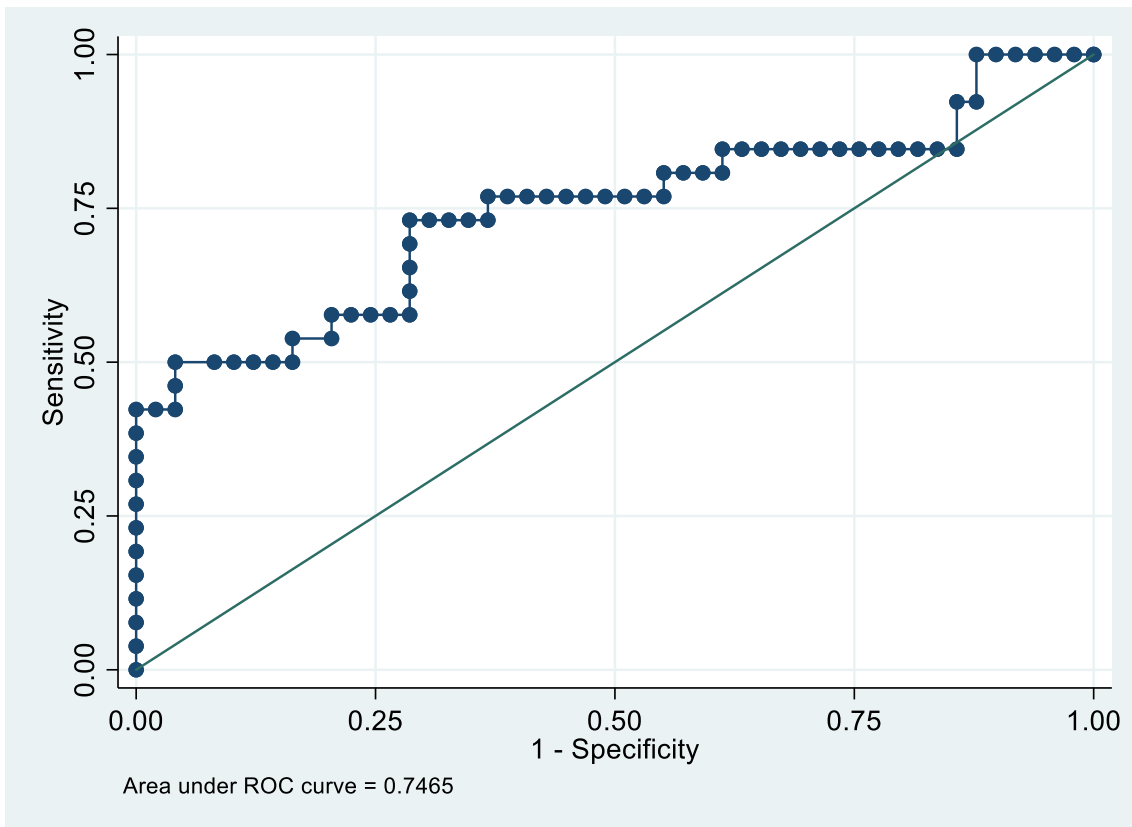
Nota: A Tabela 14 apresenta os resultados da estimação da equação (6) para uma subamostra compreendida entre 2006 e 2013, inclusive. A variável dependente é uma *dummy* que classifica as empresas com um (1) se insolvente e zero (0), caso contrário. As variáveis explicativas são as descritas detalhadamente no Subcapítulo 4.2. É, ainda, apresentado o N° de Observações, a Significância Global dos Coeficientes, o Pseudo R^2 , o Cutoff point e a percentagem de Observações Corretamente Classificadas. Os coeficientes são estatisticamente significativos ao nível de 1% (***) e 5% (**). Os erros-padrão são *White-robust*.

Da análise da Tabela 14, conclui-se que a percentagem de observações corretamente classificadas no período 2014-2021 é de 72%, sendo este valor ligeiramente superior ao obtido na Tabela 13.

Destaca-se também a Curva ROC, apresentada no Gráfico 5, com uma precisão de 74,65% na separação entre empresas insolventes e saudáveis para o período 2014-2021, sendo que na análise anterior esta precisão era de 79,83% para a amostra global. Estes resultados suportam a convicção de que o modelo final conduz a previsões com um nível de precisão aceitável (Hosmer Jr, Lemeshow, & Sturdivant, 2013).

⁷ Apesar de ser designado por teste *Out-of-Sample*, não pode ser considerado realmente assim porque as estimativas de um subperíodo serão aplicados num subperíodo subsequente da mesma amostra.

Gráfico 5 – Curva ROC do Teste *Out-of-Sample*



Nota: O Gráfico 5 representa a Curva ROC da estimação da equação (6), para a subamostra compreendida entre 2006 e 2013, inclusive, cuja estimação foi aplicada ao subperíodo 2014-2021.

7. Considerações Finais

O estudo dos modelos de previsão de insolvência tem sido desenvolvido e aplicado em inúmeros setores de atividade, em diversos países, durante largas décadas. A situação de insolvência pode derivar de fatores externos ou internos ao ambiente empresarial, sendo que é notória a elevada correlação deste evento com os ciclos económicos, bem como com problemas ao nível da gestão organizacional. Perante isto, foram vários os autores que se dedicaram a este ramo, destacando-se os trabalhos pioneiros de Beaver (1966), Altman (1968) e Ohlson (1980), muitas vezes referidos ao longo deste trabalho, seguindo-se de outros autores que adaptaram e propuseram modelos alternativos. Apesar do desenvolvimento de novos modelos mais evoluídos, o Modelo *Z-score* de Altman (1968) e o Modelo logístico de Ohlson (1980) são, ainda, os mais utilizados e referidos no estudo da previsão de insolvência financeira dada a sua fácil aplicabilidade e elevada percentagem de acerto associada.

O contexto económico e financeiro atual faz emergir a necessidade de recorrer a métodos capazes de prever uma possível situação de insolvência. A crise proveniente da pandemia da Covid-19 alarmou a sociedade e afetou a economia a nível mundial, aumentando significativamente a probabilidade de as empresas entrarem em situação de insolvência. Mais recentemente, o conflito entre a Rússia e a Ucrânia conduziu novamente à fragilidade da economia mundial, que ainda se encontrava em recuperação, uma vez que afetou a cadeia de oferta de bens e os mercados globais, resultando no aumento dos preços dos bens e serviços disponíveis, principalmente na Europa. Neste sentido, é evidenciada a importância dos modelos de previsão de insolvência, sendo que a investigação em causa tem como objetivo analisar e identificar, com recurso à utilização de rácios económico-financeiros, quais as variáveis determinantes para a previsão da insolvência das empresas cotadas da Zona Euro, através da elaboração de um Modelo *Logit*.

Para o desenvolvimento do modelo final, foi selecionada uma amostra de 981 empresas localizadas na Zona Euro e cotadas nas Bolsas de Valores Mobiliários, para um período compreendido entre 2006 e 2021. Das empresas sob análise, 51 encontram-se em estado de insolvência e as restantes em plena atividade. De forma a obter um modelo com boa capacidade preditiva, houve uma seleção prévia das variáveis que o integram, mediante o seu comportamento face a diferentes técnicas estatísticas.

Na primeira fase de seleção, as variáveis foram sujeitas a uma análise preliminar, a Análise Discriminante Múltipla. Posteriormente, os resultados obtidos nesta etapa foram complementados com os resultados da segunda fase, que representa a metodologia principal deste trabalho, a Análise Logística. Mediante as conclusões retiradas de ambas as análises, as variáveis que incorporam o modelo final de previsão de insolvência são: Rentabilidade Vendas, VM Vendas, Endividamento, Solvabilidade, Ganhos Retidos e Dimensão.

O modelo final de previsão de insolvência foi sujeito a diferentes testes de robustez, nomeadamente a análise do modelo até cinco anos antes do evento da insolvência, foi realizado o controlo para efeitos temporais e de indústria e, por fim, foi aplicado um teste *Out-of-Sample* de forma a aferir a capacidade do modelo quando aplicado numa amostra num período futuro.

O modelo final revelou uma boa capacidade na deteção da insolvência até 5 anos de, aproximadamente, 80%, o que corresponde a um valor acima do ponto de *cutoff* de 70% tal como mencionado de forma quase consensual na literatura estatística (Hosmer Jr, Lemeshow, & Sturdivant, 2013).

Apesar de o resultado final ser considerado satisfatório, este trabalho apresenta algumas limitações, nomeadamente, a falta de informação disponibilizada pela *ORBIS Europe* relativamente às empresas insolventes, dado que o número real de empresas insolventes não parece coincidir com a atribuição desta categoria disponibilizada pela base de dados, com a possibilidade de empresas insolventes estarem a ser categorizadas como saudáveis. Para colmatar esta questão, uma solução passaria por incluir no estudo empresas com rácios de cobertura inferiores a 1, ou seja, os custos de financiamento seriam superiores ao EBITDA durante um determinado período temporal consecutivo (entre dois a cinco anos), o que se traduz na incapacidade de a empresa fazer face aos seus custos de financiamento. Este comportamento constante no médio prazo conduziria à conclusão de que empresas nessa situação poderiam ser consideradas como potenciais insolventes.

Devido à conjuntura de instabilidade económica e financeira que tende a permanecer e, uma vez que são já observáveis as consequências do impacto do conflito da Rússia com a Ucrânia, as repercussões no longo prazo poderão ser severas. Assim, em investigações futuras, de forma a incrementar a capacidade preditiva dos modelos de previsão de insolvência, seria relevante incluir variáveis macroeconómicas indicadoras da qualidade institucional ao nível do país. Também a inclusão de variáveis qualitativas que permitissem medir, por exemplo, formação e qualificação

dos gestores e restantes funcionários, elementos fulcrais para um bom desempenho das organizações. Além disso, ajustar os modelos de previsão de insolvência mediante a política fiscal de cada país é igualmente um aspeto a considerar futuramente.

Referências

- Alter, C., & Hage, J. (1993). Organizations working together. *California*.
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 589-609.
- Altman, E. I. (1993). Corporate distress diagnosis: comparing using linear discriminate analysis and neural networks. *Journal of Banking and Finance*, 505-529.
- Altman, E. I., & Hotchkiss, E. (1993). *Corporate financial distress and bankruptcy: predict and avoid bankruptcy. Analyze and Invest in Distressed Debt*. New Jersey: John Wiley & Sons.
- Altman, E. I., & Sabato, G. (2013). Modeling credit risk for SMEs: evidence from the US market. *Managing and Measuring Risk: Emerging Global Standards and Regulations After the Financial Crisis*, 251-279.
- Altman, E. I., Iwanicz-Drozdowska, M., Laitinen, E. K., & Suvas, A. (2017). Financial distress prediction in an International context: A review and empirical analysis of Altman's Z-Score model. *Journal of International Financial Management & Accounting*, 131-171.
- Altman, E., Baidya, T., & Dias, L. (1979). Previsão de problemas financeiros em empresas. *Revista de Administração de Empresas*, 17-28.
- Alvares, P. M. (2019). Modelos de Previsão de Falência Empresarial: Análise Crítica do Z-score de Altman. *Universidade do Porto*, Retrieved from <https://hdl.handle.net/10216/123454>.
- Antão, M. G., Peres, C. J., & Marques, H. M. (2008). Taxonomia da Falência e a Recuperação de Empresas – O Porquê do Insucesso. *European Journal of Applied Business Management, Special Issue of ICABM2018*, 30-58.
- Antunes, F., Ribeiro, B., & Pereira, F. (2017). Probabilistic modeling and visualization for bankruptcy prediction. *Applied Soft Computing*, 831-843.
- Areal, N. M., & Armada, M. J. (1999). Testes paramétricos e não-paramétricos de reversão para a média da rentabilidade de índices do mercado accionista. *Revista de Administração Contemporânea*, 7-28.
- Banner, D. K., & Gagné, T. E. (1995). Designing effective organizations: traditional & transformational views. *Thousand Oaks, California*.
- Beaver, W. H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research*, 71-111.
- Beck, U. (2000). What Is Globalization? *Cambridge: Polity Press*.
- Blum, M. (1974). Failing company discriminant analysis. *Journal of accounting research*.
- Borooah, V. K. (2002). *Logit and Probit: Ordered and Multinomial Models. Quantitative Applications in the Social Sciences*. Thousand Oaks, California: Sage Publications.
- Cardoso, S. J. (1996). Una Evolución Crítica de la Investigación Empírica Desarrollada en Torno a la Solvencia Empresarial. *Revista Española de Financiación y Contabilidad*, 459-479.
- Casey, C., & Bartczak, N. (1985). Using operating cash flow data to predict financial distress: some extensions. *Journal of Accounting Research*, 384-401.

- Cook, R. A., & Nelson, J. L. (1998). A conspectus of business failure forecasting. *The Journal of finance*, 589-609.
- Damodaran, A. (2012). *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*. Wiley.
- Deakin, E. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 167-179.
- Diakomihalis, M. (2012). The accuracy of Altman's models in predicting hotel bankruptcy. *International Journal of Accounting and Financial Reporting*, 96-113.
- Domenichelli, O. (2018). Performance, financing decisions and corporate governance of Italian medium and large private family firms. *African Journal of Business Management*, 574-585.
- Du Jardin, P. (2010). Predicting bankruptcy using neural networks and other classification methods: The influence of variable selection techniques on model accuracy. *Neurocomputing*, 2047-2060.
- Dumontier, F. (1996). O Uso de Redes Neurais para Avaliação de Risco de Insolvência. *Revista de Administração*, 52-63.
- Eisenbeis, R. A. (1977). Pitfalls in the Application of Discriminant Analysis in Business, Finance, and Economics. *The Journal of Finance*, 875-900.
- Ferreira, P., & Caetano, J. (2013). Integração financeira incompleta: Zona Euro em cheque. *Debater a Europa* 8, 1647-6336.
- Ferreira, V. F. (2014). *Risco de Crédito*. Universidade do Porto: Faculdade de Economia do Porto.
- FitzPatrick, P. (1932). A comparison of ratios of successful industrial enterprises with those of failed companies. *The Certified Public Accountant*, 727-731.
- Galbraith, J. R. (1977). Organization design. *Addison-Wesley Publishing Company*.
- Gentry, J., Newbold, P., & Whitford, D. (1985). Classifying bankrupt firms with funds flow components. *Journal of Accounting Research*, 146-160.
- Gu, S., Bryan, K., & Dacheng, X. (2020). Empirical asset pricing via machine learning. *The Review of Financial Studies*, 2223-2273.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data mining, inference, and prediction. *Springer Science & Business Media*.
- Hatfield, G. B., Cheng, L. T., & Davidson, W. N. (1994). The determination of optimal capital structure: The effect of firm and industry debt ratios on market value. *Journal of Financial and Strategic Decisions*, 7(3), 1-14.
- Held, D., McGrew, A., Goldblatt, D., & Perraton, J. (1999). *Global Transformation*. Stanford, CA: Stanford University Press.
- Hinkle, D., Wiersm, W., & Jurs, S. (2003). *Applied Statistics for the Behavioral Sciences*, 5. Boston: Houghton Mifflin.

- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*, 398. Hoboken, New Jersey: John Wiley & Sons.
- Jackendoff, N. (1962). A study of published industry financial and operating ratios . *Philadelphia: Temple University, Bureau of economic and Business Research*.
- Jagels, M., & Coltman, M. (2004). *Hospitality management accounting*. New Jersey: John Wiley and Sons, Inc.
- Keasey, K., & Watson, R. (1987). Non-Financial Symptoms and the Prediction of Small Company Failure: a Test of Argenti's Hypotheses. *Journal of Business Finance and Accounting*.
- Lin, Y. (2014). The Relation between Auditor Switching and Self-fulfilling Prophecy Effect: The Bivariate Probit Model. *Asian Journal of Finance and Accounting*, 367-387.
- Lo, A. (1986). Logit Versus Discriminant Analysis, a Specification Test and Application to Corporate Bankruptcies. *Journal of Econometrics*, 151-178.
- Loureiro, G., & Silva, S. (2020). The impact of cross-delisting from U.S. on firms' financial constraints. *Journal of Business Research*, 132-146.
- Loureiro, G., & Silva, S. (2021). The impact of securities regulation on the information environment around stock-financed acquisitions. *Journal of International Financial Markets, Institutions and Money*, 73.
- Marín, J. L. (1977). Modelos de pronóstico de la insolvencia empresarial. In Predicción de la insolvencia empresarial. *Asociación Española de Contabilidad y Administración de Empresas, AECA*, 33-50.
- Merwin, C. (1942). Financing small corporations in five manufacturing industries. . *National Bureau of Economic Research*, 1926–1936.
- Moreira, J. A. (1998). Análise Financeira de Empresas: da teoria à prática. *Associação da Bolsa de Derivados do Porto*.
- Moreira, J. A., & Tavares, A. M. (2014). Rentabilidade do Ativo: será indiferente a definição adotada? *XVI Encontro AECA1*, 45.
- Morris, R. (1997). Early Warning Indicators of Corporate Failure – A Critical Review of Previous Research and Further Empirical Evidence. *Aldershot Ashgate*.
- Neves, J. (1997). Diagnóstico e plano de acção para recuperação - Uma visão de mercado e de crítica ao CPREF. *Seminário sobre Gestão e Liquidação Judicial: O Estado da Arte em Portugal*. .
- Neves, J. (2012). *Relato e Análise Financeira – Uma Visão Integrada de Gestão*. Alfragide: Texto Editores, Lda.
- Ohlson, J. (1980). Financial Ratios and the Probabilistic of Bankruptcy. *Journal of Accounting Research*, 109-131.
- Pampel, F. (2000). *Logistic Regression: A Primer*”, *Quantitative Applications in the Social Sciences*. Thousand Oaks: Sage Publications.
- Pinches, G. E. (1980). Factors influencing classification results from multiple discriminant analysis. *Journal of Business Research*, 429-456.

- Platt, H. D., & Platt, M. B. (1990). Development of a Class of Stable Predictive Variables: The Case of Bankruptcy Prediction. *Journal of Business Finance and Accounting*, 31-51.
- Reis, E. (2001). *Estatística Multivariada Aplicada*. Lisboa: Silabo.
- Robinson, J. A. (1998). Theories of 'Bad Policy. *The Journal of Policy Reform* 2, 1-46.
- Rodrigues Junior, M. M., Silva, T. P., & Hein, N. (2012). O uso do índice de força relativa em indicadores financeiros: Um estudo sobre a previsão de insolvência de empresas. *XXXII Encontro Nacional De Engenharia De Produção, Bento Gonçalves*.
- Sage, P. (2015). Extension chapters on advanced techniques.
- Schommer, C. P. (2000). *Empresas e Sociedade: Cooperação Organizacional num Espaço Público Comum*. Universidade Federal da Bahia - Escola de Administração.
- Sheppard, J. P. (1994). The Dilemma of Matched Pairs and Diversified Firms in Bankruptcy Prediction Models. *Mid-Atlantic Journal of Business*, 30, 9-25.
- Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *Journal of Business*, 101-124.
- Silva, A. I. (2010). *Um Modelo de Previsão de Insolvência Financeira*. Universidade da Beira Interior.
- Silva, S. S. (2015). Modelo de previsão de insolvências: uma abordagem ao setor empresarial português na indústria do calçado. *Instituto Superior de Contabilidade e Administração do Porto*.
- Smith, R. F., & Winakor, A. (1935). Changes in the Financial Structure of Unsuccessful Industrial Corporations. *Urbana: University of Illinois*.
- Sousa, J., & Oliveira, I. (2014). As variáveis de previsão de falência nas empresas portuguesas de vestuário, couro e produtos de couro. *Revista Portuguesa e Brasileira de Gestão*, 62-73.
- Sung, T. K., Chang, N., & Lee, G. (1999). Dynamics of Modeling in Data Mining: Interpretive Approach to Bankruptcy Prediction. *Journal of Management Information Systems*, 63-85.
- Taffler, R. (1983). The assessment of company solvency and performance using a statistical model. *Accounting and Business Research*, 295-307.
- Tamari, M. (1966). Financial ratios as a means of forecasting bankruptcy. *Management International Review*, 15-21.
- Zarei, S., Dadashi, I., & Akbari, M. A. (2012). Localized distress prediction models in the economic environment of Iran. *African Journal of Business Management* 6 (29), 8651.
- Zavgren, C. (1985). Assessing the vulnerability to failure of American industrial firms: a Logistic Analysis. *Journal of Business Finance & Accounting*, 12(1), 19-45.
- Zeytinoglu, E., & Akarım, Y. D. (2013). Z Financial Failure Prediction Using Financial Ratios: An Empirical Application on Istanbul Stock Exchange. *Journal of Applied Finance and Banking*, 1-8.
- Zmijewski, M. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 59-82.

- Zopounidis, C., & Doumpos, M. (1999). Business failure prediction using the UTADIS multicriteria analysis method. *Journal of the Operational research Society*, 1138- 1148.
- Zoričák, M., Gnip, P., Drotár, P., & Gazda, V. (2020). Bankruptcy prediction for small-and medium-sized companies using severely imbalanced datasets. *Economic Modelling*, 165-176. doi: 10.1016/j.econmod.2019.04.003.

Apêndice

Com o propósito de averiguar a aplicabilidade do modelo de previsão de insolvência desenvolvido em contexto real, foi selecionada, aleatoriamente, uma empresa da base de dados sob análise. A empresa selecionada, com um nome fictício, "EXEM", para o ano de 2014, apresentava os seguintes coeficientes:

- Rentabilidade Vendas = 0,0305
- VM Vendas = 0,2203
- Endividamento = 0,1450
- Solvabilidade = 2,5247
- Ganhos Retidos = 0,2637
- Dimensão = 10,7248

Sendo estes incorporados na equação (7), obtendo-se o seguinte:

$$\begin{aligned} \text{Score} &= -1,2467 + 0,1281 \times 0,0305 + 0,0243 \times 0,2203 \\ &\quad - 4,9812 \times 0,1450 - 0,0082 \times 2,5247 - 0,1651 \times 0,2637 \\ &\quad - 0,3127 \times 10,7248 = -5,3776 \end{aligned} \quad (9)$$

A partir deste resultado, é obtido o valor do *Score* que permite obter a probabilidade de a empresa entrar em insolvência, através da seguinte expressão:

$$\text{Prob}(\text{Insolvency} = 1) = \frac{1}{1 + e^{5,3776}} = 0,0046 \quad (10)$$

Considerando a zona de *cutoff* obtida para o modelo final de previsão de insolvência, de 0,4%, é possível aferir que a empresa "EXEM" é classificada como insolvente, uma vez que o seu *score* é de -5,3776, sendo este superior a 0,4%.