

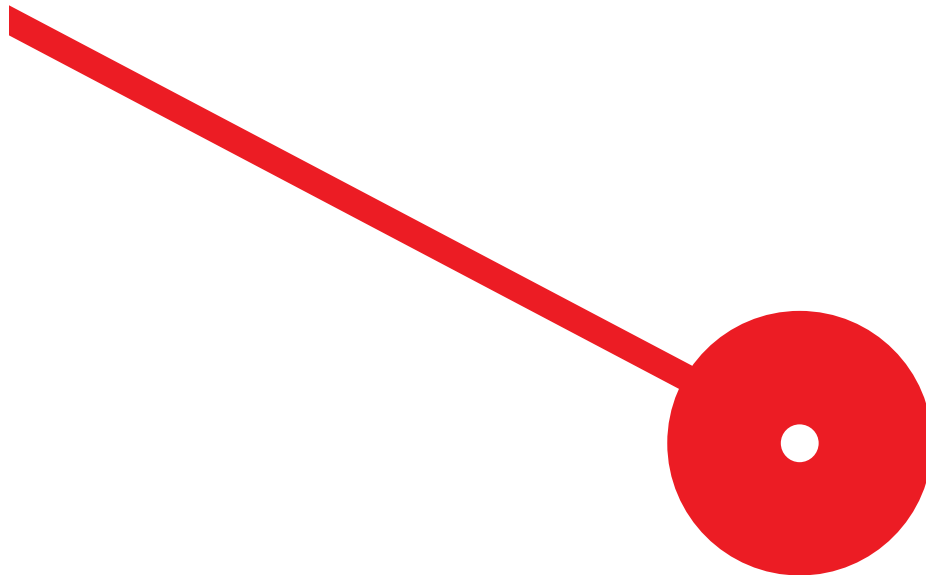


Previsão do fracasso empresarial no setor dos transportes - um estudo comparativo entre os modelos estatísticos e os modelos baseados em inteligência artificial

Mariline da Silva Lourenço

10/2023

Versão final – contém críticas e sugestões dos elementos do júri

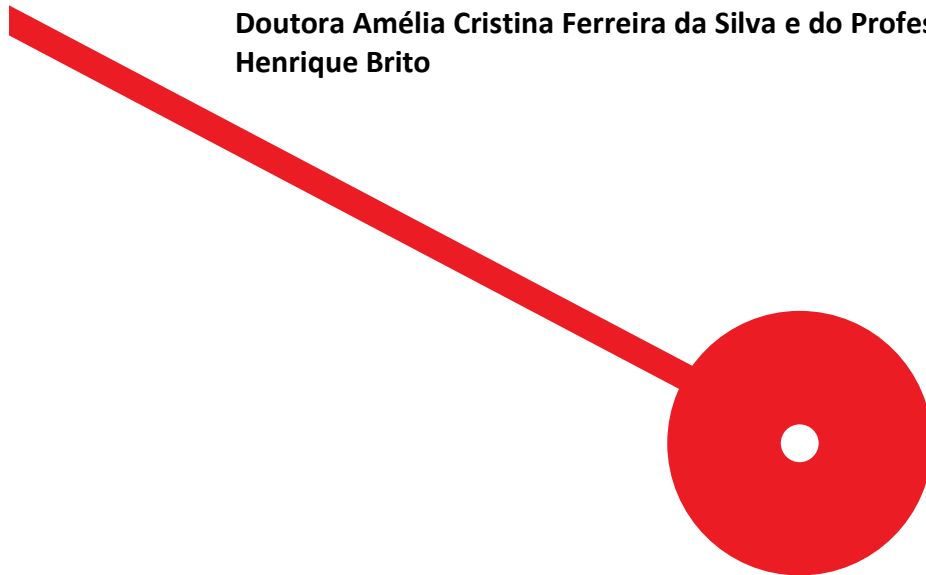




Previsão do fracasso empresarial no setor dos transportes - um estudo comparativo entre os modelos estatísticos e os modelos baseados em inteligência artificial

Mariline da Silva Lourenço

Dissertação de Mestrado apresentado ao Instituto Superior de Contabilidade e Administração do Porto para a obtenção do grau de Mestre em Contabilidade e Finanças sob orientação da Professora Doutora Amélia Cristina Ferreira da Silva e do Professor Doutor José Henrique Brito



À minha Avó Carminda.

À minha Avó Céu.

Agradecimentos

À minha orientadora, Professora Doutora Amélia Cristina Ferreira da Silva, o meu mais sincero obrigada por todo o apoio dado ao longo desta caminhada. Por desde o primeiro dia ter aceitado orientar o meu projeto, por todo o apoio na escolha do tema para esta investigação, por toda a disponibilidade, ajuda, paciência e as palavras certas no momento certo. Por querer tanto quanto eu que este projeto atingisse os resultados esperados.

Ao Professor Doutor José Henrique Brito pela ajuda imprescindível ao nível do tratamento e análise de dados, por toda a disponibilidade e apoio dado de imediato.

Aos meus pais, Anabela e Armando, obrigada por todo o apoio dado ao longo de toda a minha vida e, de modo especial, ao longo do desenvolvimento deste trabalho de investigação. Obrigada por todo o esforço que fizeram e fazem para que tenha todas as oportunidades e consiga alcançar todos os meus objetivos. Agradecer-vos nunca será o suficiente.

Ao meu irmão Afonso, que é e sempre será o meu pilar, pela sua capacidade de me fazer sorrir e de me ajudar a ultrapassar todas as dificuldades com leveza.

Aos meus avós, por todo o apoio, incentivo e motivação para que conseguisse cumprir este objetivo.

À minha afilhada que, nos seus doces 4 aninhos, tem a capacidade de não me fazer desistir nem baixar os braços.

Aos meus amigos, desde os mais antigos aos que me acompanham desde a licenciatura ou mestrado, por todo o apoio dado ao longo desta fase, por todo o companheirismo e amizade. Sem vocês isto não seria possível!

A todos que se cruzaram na minha caminhada, o meu mais sincero obrigada!

Resumo

O setor dos transportes é um setor crucial e bastante importante para o nosso dia-a-dia, na medida em que é o único que permite transportar mercadorias e pessoas entre os diversos países. É, portanto, um setor insubstituível e indispensável sendo uma parte essencial do tecido funcional da nossa sociedade.

O insucesso empresarial, por sua vez, é um fenómeno que apresenta bastante interesse para qualquer ramo, incluindo o ramo dos transportes. Tal facto tem levado a que sejam desenvolvidos modelos de previsão do fracasso empresarial que, a partir de de um conjunto de indicadores, indicam a probabilidade de a organização falhar.

Aliada a esta situação, assistimos diariamente a uma permanente evolução tecnológica, bem como, a uma constante busca por métodos mais eficazes na na previsão do fracasso empresarial, levando os investigadores a recorrer, numa tendência crescente, à inteligência artificial.

O presente trabalho visa efetuar um estudo comparativo entre os modelos de fracasso empresarial estatísticos e os modelos baseados em inteligência artificial para o setor dos transportes, num horizonte temporal de 2014 a 2021 e para os países de Portugal, Espanha, França e Itália, de modo a verificar qual dos modelos é mais eficiente. A amostra utilizada foi constituída por um conjunto final de 4866 empresas, das quais 2881 são não fracassadas e 1985 fracassadas.

Os modelos desenvolvidos permitiram classificar corretamente entre 71% a 73% das empresas. Não obstante, quando estes resultados são comparados com os resultados obtidos nos modelos de previsão do fracasso empresarial estatísticos , constata-se que estes últimos apresentam uma capacidade preditiva superior, registando menos erros na classificação das empresas em análise

Palavras chave: Inteligência Artificial; Fracasso Empresarial; Modelos de Previsão; Indicadores Financeiros; Setor Transportes.

Abstract

The transport sector is crucial and very important to our daily lives, as it is the only sector that allows us to transport goods and people between different countries. It is therefore an irreplaceable and indispensable sector and an essential part of the functional fabric of our society.

Business failure, on the other hand, is a phenomenon that is of great interest to any industry, including transport. This has led to the development of business failure prediction models which, based on a set of indicators, indicate the likelihood of the organization failing.

Allied to this situation, we are witnessing constant technological evolution on a daily basis, as well as a constant search for new, more effective methods that are capable of predicting business failure, leading to a growing trend towards the implementation of artificial intelligence to estimate models for predicting business failure.

The aim of this paper is to carry out a comparative study between statistical business failure models and models based on artificial intelligence for the transport sector, over a time horizon of 2014 to 2021 and for the countries of Portugal, Spain, France and Italy, in order to see which of the models is more efficient. The sample used consisted of a final set of 4866 companies, of which 2881 were non-failed and 1985 failed.

The results of this study made it possible to correctly classify between 71% and 73% of the companies, as well as to verify that when statistical business failure prediction models are compared with prediction models based on artificial intelligence, the latter have a higher predictive capacity, registering fewer errors in the classification of the companies under analysis as failed or non-failed.

Keywords: Artificial Intelligence; Business Failure; Forecasting Models; Financial Indicators; Transportation Sector.

Índice geral

Capítulo - Introdução	12
Capítulo I – Revisão de Literatura	16
1. Revisão de Literatura.....	17
1.1 Conceito de Fracasso.....	17
1.2 Modelos de Previsão do Fracasso Empresarial	18
1.2.1 Modelo Estatísticos.....	20
1.2.1.1 Modelo de Beaver (1966) – Análise Univariada	20
1.2.1.2 Modelo de Altman (1968) – Análise Discriminante Múltipla (ADM)	21
1.2.1.3 Modelo de Ohlson (1980) – Modelo <i>Logit</i>	23
1.2.2 Modelos baseados em Inteligência Artificial.....	24
1.2.2.1 Redes Neurais (NN).....	25
1.2.2.2 Árvores de Decisão (DT).....	27
1.2.2.3 Máquinas Vetoriais de Suporte (SVM)	29
1.2.2.4 Algoritmo Genético (AG).....	30
Capítulo II – O Setor dos Transportes	33
2 O Setor dos Transportes	34
2.1 Situação Atual do Setor.....	34
2.2 Tipos de Transportes	35
Capítulo III – Metodologia	38
3 Metodologia.....	39
3.1 Seleção das variáveis.....	39
3.2 Amostra e conjunto de dados	40
3.3 Preparação dos dados	41
Capítulo IV – Modelos e Análise dos Resultados	43
4 Modelos e Análise dos Resultados	44

4.1	Modelos Utilizados	44
4.2	Resultados e Análise	44
Capítulo	– Conclusão	47
Referências bibliográficas	51

Índice de Figuras

Figura 1: Palavras-chave baseadas na ocorrência.....	18
Figura 2: Cooperação entre países	19
Figura 3: Rede de autores baseada na coautoria dos autores principais	19
Figura 4: Exemplo de uma rede neural.....	27
Figura 5: Exemplo de uma árvore de decisão.....	28

Índice de Tabelas

Tabela 1: Resultados dos modelos por cada tipo de classificador.....	45
----------------------------------------------------------------------	----

Lista de siglas

ADM – Análise Múltipla Discriminante

AG – Algoritmo Genético

CAE – Código da Atividade Económica

DT – Árvores de Decisão (Decision Trees)

ERF – *Extremely Randomized Forest*

KNN – *K-Nearest Neighbors*

LR – Regressão Linear

NN – Redes Neurais (*Neural Networks*)

PIB – Produto Interno Bruto

RF – *Random Forests*

SVM – Máquina Vetorial de Suporte (*Support Vector Machines*)

UE – União Europeia

A previsão do fracasso empresarial é uma temática bastante importante e que tem revelado um crescente interesse nas últimas décadas, isto porque, seguindo a linha de pensamento de Borchert et al. (2023), prever o insucesso empresarial fornece aos investigadores e colaboradores um instrumento muito relevante para avaliar a saúde financeira da empresa. Aliada a esta situação e em concordância com Pereira et al. (2012), a saúde financeira das empresas tem sido uma das maiores preocupações sociais para os agentes económicos pelo que existe uma vasta gama de modelos para prever o insucesso empresarial utilizados para discriminar entre empresas fracassadas e empresas não fracassadas. É, de facto, um assunto que deve ser constantemente estudado e adaptado à realidade económica envolvente de forma a se obterem resultados cada vez mais precisos.

Numa breve definição, fracasso pode ser entendido como a incapacidade de a empresa conseguir fazer face a todas as suas obrigações financeiras levando, posteriormente, a uma possível falência. Os modelos de previsão do fracasso empresarial são desenvolvidos no sentido de orientar as empresas para uma possível situação de fracasso, através de um conjunto de indicadores, ajudando as mesmas a tomarem medidas preventivas.

Têm, portanto, sido realizados vários estudos com vista ao desenvolvimento de novos modelos de previsão do fracasso empresarial, bem como a melhorar a capacidade preditiva de cada modelo. O estudo de Beaver (1966) destaca-se por ser o primeiro estudo realizado na temática dos modelos de previsão do fracasso empresarial utilizando rácios financeiros para estimar o modelo. Posteriormente, foram desenvolvidos novos modelos como é o caso do modelo de análise discriminante múltipla ((Altman (1968), Deakin (1972)), do modelo logit (Ohlson (1980)) e probit ((Zmijewski, 1984)) e, mais recentemente, os modelos baseados em inteligência artificial de que são exemplos as redes neurais (Altman et al. (1994), Neves & Vieira (2006)), as árvores de decisão (Gepp et al. (2010), Pereira et al. (2010)) e as máquinas vetoriais de suporte (Alaka et al. (2018), Shetty et al. (2022)), assim como o algoritmo genético (Gordini, 2014).

Por sua vez, o setor dos transportes é um setor muito importante e imprescindível à necessidade humana, permitindo transportar pessoas e bens entre países. As empresas do ramo dos transportes de mercadorias apresentam um peso significativo tanto no tecido empresarial português, como no tecido empresarial espanhol, como no tecido empresarial francês e também no tecido empresarial italiano. De acordo com dados do Banco de Portugal, em 2021 havia em Portugal 21 497 empresas do ramo dos transportes incluindo

tal número empresas de transportes rodoviários, empresas de transportes marítimos e empresas de transportes aéreos.

Deste modo, sendo o setor dos transportes muito relevante, este trabalho de investigação visa efetuar uma comparação entre os modelos de fracasso empresarial estatísticos e os modelos baseados em inteligência artificial no setor dos transportes, num intervalo de tempo compreendido entre 2014 e 2021 e para os países de Portugal, Espanha, França e Itália. A escolha deste horizonte temporal deve-se ao facto de o surgimento da pandemia da COVID-19 em 2020 ter condicionado todos os negócios levando a uma quebra no volume de negócios e lucros das empresas. Deste modo, utilizando este horizonte temporal será também possível verificar o efeito da pandemia numa eventual situação de fracasso das organizações. Já relativamente aos países em análise, optei por esta escolha pelo facto de serem dos principais países para a realização de trocas comerciais por meio dos transportes e também por serem aqueles com maior tráfego de transporte de mercadorias entre si.

O principal objetivo desta investigação consiste em comparar o desempenho dos modelos de previsão do fracasso empresarial estatísticos e os modelos de fracasso empresarial baseados em inteligência artificial no setor dos transportes, visando, com isto, analisar qual dos modelos se torna mais eficiente. Seguindo a linha de pensamento de J. Pereira et al. (2014), as técnicas de inteligência artificial usadas para prever o fracasso empresarial são mais eficientes e apresentam uma maior taxa de assertividade e, portanto, pretendo desenvolver este trabalho no sentido de verificar se, de facto, os modelos baseados em inteligência artificial são mais precisos.

Sendo a previsão de sucesso empresarial um fator importante para as organizações e que ganhou especial interesse a partir da crise financeira global de 2008, de acordo com o reportado por Shi & Li (2019), pode-se verificar a pertinência deste estudo, bem como se pode referir que o mesmo permitirá aprofundar alguns estudos já efetuados nesta temática.

Relativamente às questões de investigação, pretendo responder às seguintes:

- i) Quais os fatores que influenciam o fracasso empresarial no setor dos transportes?
- ii) Quais os indicadores mais relevantes para a previsão do fracasso empresarial?

- iii) Os modelos de previsão do fracasso empresarial baseados em inteligência artificial apresentam um melhor desempenho que os modelos estatísticos?

No desenvolvimento desta investigação foi utilizada uma amostra composta por um total de 4866 empresas e um conjunto de 15 rácios financeiros utilizados para prever o modelo. Todos os dados usados foram recolhidos com recurso à base de dados Bureau Van Dijk – ORBIS.

A presente dissertação encontra-se dividida em seis capítulos. Após a introdução, no capítulo um será apresentada uma revisão de literatura acerca do conceito de fracasso e também dos modelos de fracasso empresarial, fazendo-se uma breve referência a alguns dos modelos mais comuns na literatura. De seguida, o capítulo dois incidirá sobre um breve enquadramento do setor em estudo com uma curta referência à situação atual do mesmo. No capítulo três será apresentada a metodologia utilizada e a descrição da amostra que serve de base para este estudo. No capítulo quatro será desenvolvido o estudo empírico desta investigação, assim como será realizada a respetiva análise e discussão dos resultados obtidos. Por fim, serão referidas as principais conclusões desta investigação, os principais entraves e propostas para investigações futuras.

CAPÍTULO I – REVISÃO DE LITERATURA

1. Revisão de Literatura

1.1 Conceito de Fracasso

Existe um vasto conjunto de definições associadas ao conceito de fracasso e fracasso empresarial dada a constante preocupação em prever a probabilidade de insucesso de qualquer organização. Tendo em conta o referido por Walsh & Cunningham (2016), as primeiras investigações sobre o fracasso empresarial começaram a surgir no século XIX tendo-se verificado uma investigação mais intensiva e aprofundada nesta temática nos últimos anos.

De acordo com Beaver (1966), fracasso pode ser definido como a incapacidade de a organização conseguir pagar todas as suas obrigações financeiras, isto é, quando se verifica a falta de pagamento aos credores, o não pagamento de dividendos aos acionistas ou quando se verifica falência.

À semelhança do referido anteriormente, também Beaver et al. (2010) definem fracasso como o incumprimento de obrigações financeiras. Neste caso, os autores associam o termo de fracasso ao termo de dificuldades financeiras que engloba dois eventos principais: por um lado temos o incumprimento das obrigações quando devidas e, por outro lado, temos a falência da organização.

Por outro lado, seguindo a linha de pensamento de Altman & Hotchkiss (2006) podemos afirmar que uma dada empresa pode encontrar-se numa situação de fracasso económico durante determinado tempo sem nunca deixar de cumprir com as suas obrigações, sendo, para estes autores, a própria qualidade da gestão por parte da administração da entidade o principal motivo para a ocorrência de fracasso.

Já em concordância com Amankwah-Amoah et al. (2022), o fracasso empresarial acontece quando uma empresa é incapaz de “dar a volta” após uma situação de declínio o que conduz ao seu colapso. Em outro ponto e de forma mais geral, o fracasso empresarial pode ser definido como a situação em que a empresa cessa as suas operações devido à incapacidade de responder e de se adaptar às mudanças externas (Amankwah-Amoah & Wang, 2019).

Importa, contudo, referir que em conformidade com J. M. Pereira et al. (2007) não existe um conceito único de fracasso empresarial, podendo este assumir a forma de falência jurídica, insolvência, suspensão de pagamentos ou, até mesmo, perdas contínuas.

1.2 Modelos de Previsão do Fracasso Empresarial

Idêntico àquilo que se verifica em relação ao conceito de fracasso, também no que toca aos modelos de previsão do fracasso empresarial encontramos um vasto conjunto de conceitos e de modelos que podem ser usados, desde o modelo pioneiro de análise univariada de Beaver (1966) à análise discriminante multivariada, ao modelo logit e probit, de acordo com J. M. Pereira et al. (2007), e mais recentemente aos modelos baseados em inteligência artificial.

Em concordância com Korol & Spyridou (2020) é essencial que numa organização exista um sistema de alerta financeiro que use um modelo de previsão razoável permitindo aos intervenientes analisar os riscos financeiros que podem determinar o sucesso ou fracasso da organização.

Importa inicialmente referir que todos os modelos abordados apresentam as suas vantagens e desvantagens, não existindo por isso uma técnica ótima e levando a que cada um adote o seu próprio modelo e conceito de fracasso empresarial J. Pereira et al. (2014).

Com recurso ao VOSviewer, numa pesquisa pelos artigos seleccionadas as principais palavras-chave encontradas estão presentes na figura 1, com principal destaque para “fracasso empresarial”, “previsão do fracasso empresarial” e “rácios financeiros” que tomaram especial atenção nesta revisão.

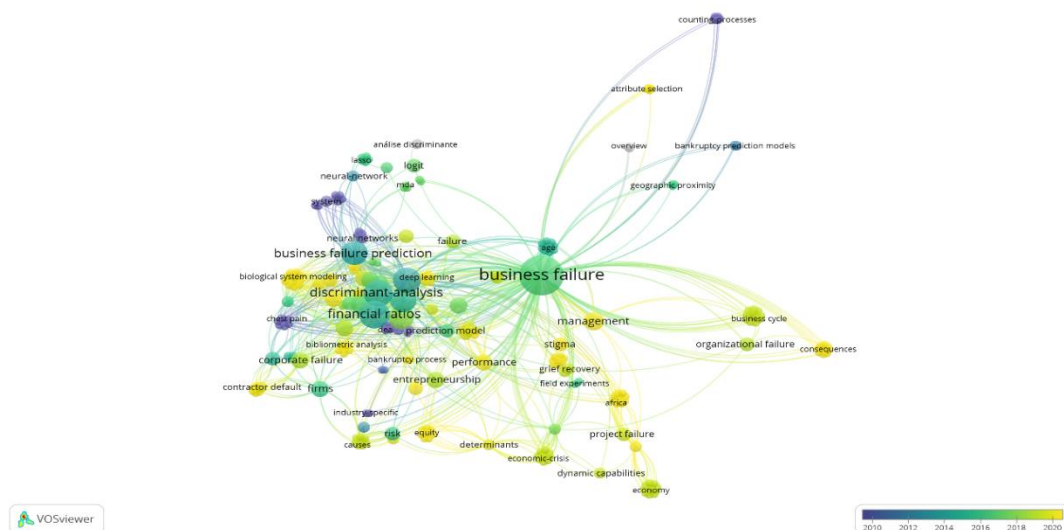


Figura 1: Palavras-chave baseadas na ocorrência
Fonte: [VOSviewer]

Em relação à cooperação entre os países que publicaram artigos acerca do fracasso empresarial, a figura 2 mostra os principais 12 países dos artigos abordados neste estudo. Destaque para os Estados Unidos da América, representado pelo cluster a azul, que apresentam o país com o maior número de publicações e de citações, seguido da Espanha no cluster a amarelo e da Inglaterra no cluster a vermelho.



Figura 2: Cooperação entre países
Fonte: [VOSviewer]

Em termos de autores com o maior número de publicações nos artigos analisados podemos destacar os estudos de Edward Altman, William Beaver e também James Ohlson. Já em termos de autores citados, a figura 3 mostra-nos uma rede colaborativa baseada na coautoria dos autores principais, sendo que neste caso apenas 8 autores estavam conectados entre si, com destaque para o cluster representado a verde.

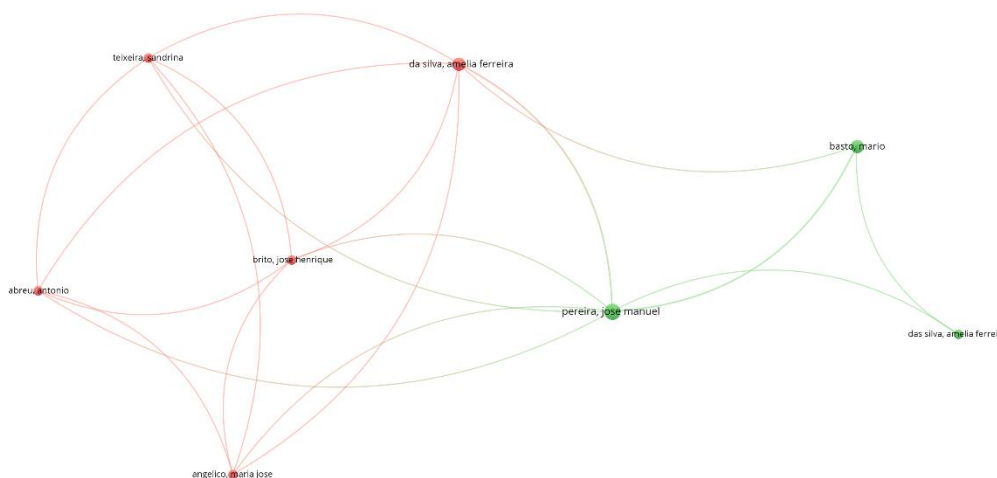


Figura 3: Rede de autores baseada na coautoria dos autores principais
Fonte: [VOSviewer]

Em concordância com J. M. Pereira et al. (2016) os modelos de fracasso empresarial começaram a surgir na década de sessenta dada a necessidade de se analisar o insucesso empresarial e as respetivas consequências económicas e sociais para os envolvidos. Ora, segundo os mesmos autores, as primeiras técnicas utilizadas passavam por técnicas estatísticas aplicadas aos dados contabilísticos das empresas com recurso a um conjunto

de rácios financeiros. Estudos de Beaver (1966) e Altman (1968) são dos primeiros a relatar a utilização de rácios financeiros para prever o insucesso empresarial e também são os principais estudos citados nesta análise. Em termos de técnica mais utilizada para prever o fracasso empresarial, seguindo a linha de pensamento de Alaka et al. (2018) verifica-se que a análise discriminante múltipla é a mais comum e presente nos estudos em análise.

1.2.1 Modelo Estatísticos

1.2.1.1 Modelo de Beaver (1966) – Análise Univariada

Segundo Beaver (1966), autor do primeiro modelo de previsão do fracasso empresarial baseado em métodos estatísticos, prever o fracasso empresarial de uma organização passa por realizar uma análise univariada dos rácios financeiros de forma individual, sem ter em conta eventuais relações que existam entre eles. Este modelo assenta, portanto, numa classificação dicotómica das empresas – empresa fracassada ou não fracassada (Maricica & Georgeta, 2012).

Assim, este modelo que, segundo Shetty et al. (2022), introduziu a primeira justificação estatística para a capacidade de os rácios financeiros contabilizarem o incumprimento, é feito por três vias:

- (1) Comparação de médias;
- (2) Teste de classificação dicotómico;
- (3) Análise dos rácios de probabilidade.

Inicialmente é feita uma comparação do valor da média dos rácios em análise, no caso Cash-flow/Passivo Total, Resultado Líquido/Volume de negócios, Passivo Total/Ativo Total, Fundo de Maneio/Ativo Total, Liquidez Geral e *No-credit interval*, de cada uma das empresas, fracassadas e não fracassadas. De seguida, os rácios são dispostos numa matriz e analisados tendo por base um ponto ótimo que irá definir se a empresa é fracassada ou não. Por fim, são comparados os dados obtidos neste teste com os resultados reais da amostra de modo a se obter a taxa de erro do modelo. Trata-se de um modelo bastante eficaz com a capacidade de classificar corretamente 78% das empresas até cinco anos antes do seu insucesso (Beaver, 1966).

No entanto, segundo o próprio autor este modelo apresenta algumas limitações na medida em que a utilização de apenas uma variável independente pode conduzir a resultados errados e contrários à realidade.

1.2.1.2 Modelo de Altman (1968) – Análise Discriminante Múltipla (ADM)

Seguindo a linha de pensamento de J. M. Pereira et al. (2010), a análise discriminante surgiu com vista a distinguir estatisticamente dois ou mais grupos de organizações em relação a um conjunto de variáveis em simultâneo, coisa que não acontecia no modelo de análise univariada. Neste caso e à semelhança do que se verificava no modelo previsto por Beaver (1966) também no modelo de Altman (1968) são usados rácios financeiros para estimar o fracasso empresarial, sendo neste caso utilizada uma análise discriminante múltipla (ADM) de modo a combinar possíveis relações que existam entre os rácios.

Em concordância com J. Pereira et al. (2014) o modelo de Altman é considerado o modelo clássico para a previsão de falências, sendo também o mais usado e mais referido na literatura.

Altman (1968) apresenta um modelo denominado por modelo Z-score com uma dimensão de análise mais reduzido uma vez que a variável independente, no caso uma variável qualitativa, apenas pode tomar duas possibilidades – empresa fracassada ou não fracassada. O principal objetivo deste modelo passa por se obter uma pontuação de “Z” resultante da combinação linear de vários fatores independentes (J. Pereira et al., 2014).

A análise discriminante múltipla baseia-se na função discriminante apresentada mais abaixo e cujo cálculo nos dá apenas um valor a partir do qual conseguiremos classificar a empresa como fracassada ou não.

$$Z = V_i X_j + V_{ii} X_{jj} + \dots + V_n X_n$$

onde X_j representam as variáveis independentes – rácios económicos e financeiros, nomeadamente Fundo de maneo/Ativo total, Resultados transitados/Ativo total, EBIT/Ativo total, Capitalização bolsista/Passivo total e Vendas/Ativo total, V_{ii} os valores dos coeficientes discriminantes calculados através da ADM e Z a variável dependente.

Após obter o modelo foram testadas cada uma das empresas da amostra e comparados os resultados obtidos pelo modelo Z-score com a situação real das empresas obtendo-se,

assim, a taxa de erro deste modelo (Altman, 1968). Conclui-se assim que, segundo o modelo de ADM e em concordância com o referido por Lennox (1999), uma empresa terá mais probabilidades de fracassar se não for rentável, se for demasiado alavancada e ainda se sofrer de dificuldades financeiras.

Deste modo e em concordância com Altman (1968), tendo por base as empresas que foram mal classificadas através deste modelo, assume-se que as empresas que apresentem um valor $Z > 2,99$ sejam classificadas como não fracassadas enquanto que aquelas que registem um valor de $Z < 1,81$ são consideradas como fracassadas. As empresas com $1,81 < Z < 2,99$ são consideradas na “zona de ignorância” para se evitar erros de classificação.

Em suma, Altman (1968) verificou que o modelo de análise discriminante múltipla é bastante preciso uma vez que conseguiu prever corretamente 94% das empresas utilizadas na amostra. A juntar a isso, J. M. Pereira et al. (2016) referem como uma das vantagens do modelo a “capacidade para incorporar múltiplos rácios financeiros simultaneamente”, assim como ser de aplicação bastante simples. J. M. Pereira et al. (2017) testaram também o modelo de ADM no seu estudo e verificaram que o mesmo tem uma taxa de acertividade bastante alta, no caso de 96%.

Contudo, apresenta algumas limitações como por exemplo o facto de o modelo não incluir dados qualitativos, ou seja, eventuais acontecimentos que tenham afetado a atividade da empresa e que não estejam representados nas demonstrações financeiras. Ohlson (1980) acrescenta como limitação da ADM o facto de o resultado obtido ser de pouca interpretação intuitiva dado que se trata apenas de uma classificação ordinal, bem como a classificação das empresas como falhadas ou não falhadas ser feita de acordo com critérios como a dimensão da mesma ou a indústria, que tendem a ser arbitrários.

Alguns estudos foram realizados tendo em vista comparar qual dos dois modelos – análise univariada ou ADM – era o mais eficiente, como é o caso do estudo de Deakin (1972). O mesmo verificou que o modelo de análise univariada mostra-se mais eficiente para o curto prazo enquanto que o modelo de ADM se mostrou mais eficaz no longo prazo, através da comparação das taxas de erro.

1.2.1.3 Modelo de Ohlson (1980) – Modelo *Logit*

Com todas as limitações apontadas aos modelos pioneiros para prever o fracasso empresarial surge a necessidade de se experimentar novas técnicas que sejam mais precisas e que apresentem uma taxa de erro mais reduzida (J. Pereira et al., 2014).

Deste modo, numa abordagem significativamente diferente da até então apresentada, Ohlson (1980) desenvolveu um modelo de previsão do fracasso empresarial baseado numa análise de regressão logística, nomeadamente uma análise condicional *logit*, modelo esse que revelou um crescimento significativo da sua utilização a partir da década de oitenta (Hoetker, 2007). Segundo Ohlson (1980), o seu principal objetivo ao utilizar este novo modelo de previsão do fracasso empresarial passa por ultrapassar todas as dificuldades da ADM, conseguindo além de classificar a empresa como fracassada ou não, calcular a probabilidade de a mesma vir a entrar em insolvência num período próximo.

Assim, o modelo *logit* para a previsão do fracasso empresarial assenta na seguinte expressão:

$$P(X_i, \beta) = \frac{1}{1+e^{(-y_i)}} \text{ e } y_i = \sum \beta_j X_{ij}$$

onde P representa a probabilidade de falência da empresa, assumindo valores compreendidos entre 0 e 1 e y_i é o resultado da soma de todas as variáveis independentes usadas – rácios económicos e financeiros, com destaque para a utilização de uma medida de performance (Ohlson, 1980).

Hauser & Booth (2022) apontam alguns inconvenientes da análise *logit* para a previsão do fracasso empresarial com especial destaque para a interpretação da percentagem do modelo classificado corretamente, dado que a mesma pode ser enganadora quando os rácios utilizados apresentam resultados bastante altos. J. M. Pereira et al. (2016) acrescentam como limitação o facto de o modelo não considerar probabilidades anteriores bem como os erros de classificação.

Semelhante ao modelo desenvolvido por Ohlson existe o modelo *probit* que, seguindo a linha de pensamento de Hoetker (2007), trata-se de um modelo que deve ser utilizado quando se pretende saber qual de duas hipóteses ocorre, no caso, a empresa é considerada fracassada ou não fracassada. Importa referir que o modelo *probit*, em concordância com

J. Pereira et al. (2014), não é muito popular neste ramo, não sendo por isso tão usado como os restantes modelos.

Estudos que comparam a eficiência dos modelos *logit* e *probit* verificaram que não existe praticamente nenhuma diferença entre o poder explicativo de um ou de outro modelo, apresentando uma capacidade preditiva bastante semelhante. No entanto, os resultados obtidos através do modelo *probit* são mais difíceis de interpretar, o que explica a menor utilização deste modelo (J. M. Pereira et al., 2007). Quando comparados estes modelos com o modelo de análise discriminante múltipla, à semelhança do que acontece no estudo de Lennox (1999), verifica-se que os modelos *logit* e *probit* apresentam um melhor desempenho.

1.2.2 Modelos baseados em Inteligência Artificial

As técnicas estatísticas são, portanto, as mais usadas para previsão do fracasso empresarial, escolhendo, com base num ponto de corte previamente definido, se a empresa é classificada como fracassada, caso apresente uma pontuação inferior ao ponto ótimo, ou não fracassada, caso contrário.

Contudo, em concordância com Ooghe & Spaenjers (2009) podem ser encontrados alguns erros nestas técnicas, erros esses que acarretam alguns custos para as empresas, nomeadamente erros do tipo I quando uma empresa fracassada é classificada como não fracassada, ou erros do tipo II quando se verifica o contrário.

J. M. Pereira et al. (2016) acrescentam ainda que os primeiros trabalhos apenas tinham como objetivo provar se os dados continham informação suficiente acerca de uma futura insolvência ou então apenas comparar os modelos de modo a se descobrirem quais as melhores alternativas.

Yeh et al. (2010) apontam como limitações dos primeiros modelos de previsão do fracasso empresarial o facto de os mesmos apenas utilizarem rácios financeiros como variáveis independentes e, em contrapartida, o facto de não considerarem a eficiência da gestão da própria empresa como uma variável chave para a classificação das mesmas como fracassada ou não. A juntar a isso acresce o facto de as técnicas estatísticas apresentarem alguns pressupostos restritivos como a linearidade, normalidade e independência entre as variáveis em estudo (W. W. Wu, 2010).

Ora, atendendo a que tal violação destes pressupostos pode por em causa a obtenção de dados eficazes e válidos, seguindo a linha de pensamento de K.-S. Shin & Lee (2002), surge a necessidade de se encontrarem novas técnicas para avaliar a previsão do fracasso que sejam mais precisas e que apresentem menos erros, de que é exemplo as técnicas de previsão do fracasso empresarial baseadas em inteligência artificial.

De acordo com o mencionado por Shetty et al. (2022), os anos 90 marcam a segunda fase da história dos modelos de previsão do fracasso empresarial com a introdução de novos métodos, nomeadamente algoritmos de inteligência artificial como são exemplos as redes neurais e as árvores de decisão. Seguindo a linha de pensamento de Yeh et al. (2010), as técnicas para a previsão do fracasso empresarial baseadas em inteligência artificial podem ser alternativas aos métodos estatísticos tradicionais, permitindo contornar os principais problemas dos mesmos.

A grande diferença entre as técnicas de previsão estatísticas e as baseadas em inteligência artificial prende-se com o facto de os pressupostos e os algoritmos serem usados para gerar estruturas de conhecimento (K. S. Shin & Lee, 2002).

1.2.2.1 Redes Neurais (NN)

As redes neurais, ou *neural networks*, são das técnicas baseadas em inteligência artificial mais utilizadas para prever o fracasso empresarial e constituem uma alternativa aos modelos de previsão estatísticos, uma vez que, segundo Dong & Chen (2009) as mesmas apresentam uma capacidade de previsão significativamente superior à dos modelos estatísticos.

Neves & Vieira (2006) definem redes neurais como “um conjunto de algoritmos inspirados na arquitetura do cérebro humano e nas capacidades de processamento em paralelo”, isto é, um conjunto de máquinas que são capazes de aprender diretamente a partir de um exemplo e que não necessitam de nenhum conhecimento prévio de qualquer problema. Uma rede neural é constituída por um conjunto de unidades de processamento que são homogéneas e estão interligadas entre si, sendo que cada unidade tem uma função de cálculo e cujo comportamento é modelado por uma função matemática (Tam & Kiang, 1992).

Uma das funções mais utilizadas neste tipo de modelos, representada mais abaixo, segundo Altman et al. (1994) permite obter uma resposta – empresa fracassada ou não fracassada – sendo posteriormente comparada com o valor real. Caso se obtenham resultados diferentes será repetido o processo até se reduzir o nível de erro a mínimos aceitáveis e aí se considera que a rede alcançou o resultado de equilíbrio, ou seja, foi capaz de resolver o problema.

$$Y = \frac{1}{1 + e^{-Pi}}$$

As redes neurais apresentam um vasto conjunto de vantagens quando comparadas com os modelos estatísticos e, em concordância com Neves & Vieira (2006), a principal vantagem das mesmas prende-se com o facto de não “obrigarem” a que exista uma relação funcional entre as variáveis em estudo, uma vez que o conhecimento está diretamente ligado por se tratarem de máquinas de aprendizagem. Altman et al. (1994) acrescentam como vantagem das redes neurais o facto de os resultados obtidos através deste método derivar de um comportamento coletivo de várias unidades e não apenas de uma unidade em específico. Ao mesmo tempo e à semelhança do que se verifica nos modelos estatísticos, seguindo a linha de pensamento de Gámez et al. (2016), as mesmas são de aplicação bastante fácil.

Contudo, este método de previsão apresenta alguns contras, de que são exemplos o facto de se tratar de um processo de aprendizagem demasiado lento, o sistema pode não atingir um nível de erro minimamente aceitável e também a análise dos resultados ser bastante complexa e de difícil interpretação (Altman et al., 1994).

Podemos ainda acrescentar, em concordância com K.-S. Shin & Lee (2002), como desvantagem do modelo o facto de ser difícil para o utilizador de compreender facilmente as regras finais que estes modelos adquirem.

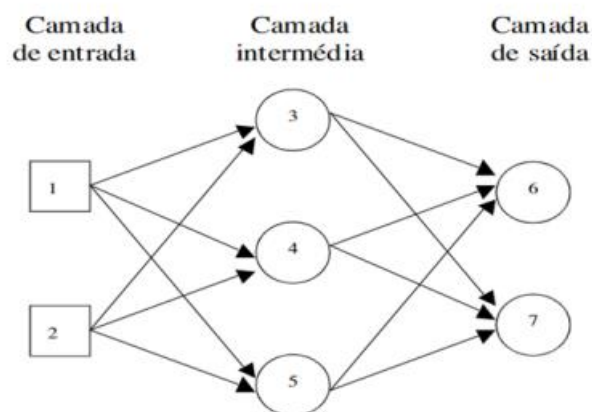


Figura 4: Exemplo de uma rede neural
 Fonte: [J. M. Pereira et al. (2007)]

Existem alguns estudos que comparam a utilização de técnicas estatísticas com a utilização de redes neurais, como é o caso do estudo de Altman et al. (1994) que realizou a comparação entre a ADM e a análise *logit* com as redes neurais. Os mesmos autores verificaram que ambas as técnicas são bastante precisas ainda que, no entanto, em algumas situações as redes neurais apresentem uma capacidade preditiva superior às técnicas estatísticas. Também Gámez et al. (2016) verificaram que a capacidade preditiva das redes neurais é superior à da ADM e também à do modelo *logit* e *probit*.

Altman et al. (1994) defendem ainda que, em estudos futuros, será adequado combinar a utilização de redes neurais e da análise discriminante múltipla de forma a se obterem resultados mais exatos, mais lógicos e de maior facilidade de compreensão.

1.2.2.2 Árvores de Decisão (DT)

As árvores de decisão são, à semelhança das redes neurais, um conjunto de técnicas de aprendizagem automática que podem ser utilizadas para prever o fracasso empresarial, não sendo ainda tão comuns em estudos realizados por serem um método relativamente recente (J. M. Pereira et al., 2007).

De acordo com J. M. Pereira et al. (2007), as árvores de decisão consistem numa forma de representação de um dado conjunto de regras seguidas por uma hierarquia de classes ou valores tendo por base uma lógica simples condicional. O termo “árvores” deve-se ao facto de, em termos gráficos, serem semelhantes a uma árvore através de uma estrutura que interliga um conjunto de nós terminais que representam os grupos de

classificação empresa fracassada ou não fracassada – e um nó raiz que, por sua vez, representa a regra de divisão ou decisão (Gepp et al., 2010).

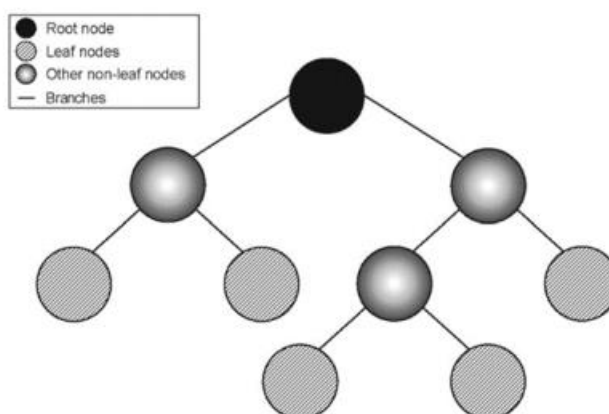


Figura 5: Exemplo de uma árvore de decisão
Fonte: [Gepp et al. (2010)]

Seguindo a linha de pensamento de J. M. Pereira et al. (2010), o principal objetivo da utilização de árvores de decisão para prever o fracasso empresarial passa por encontrar possíveis relações e dependências entre as variáveis utilizadas sendo, geralmente, os resultados apresentados sob a forma de regras, por exemplo “ $X \rightarrow Y$ ” que significa que “se X estiver presente, então é provável que Y também esteja presente”, representando X e Y as variáveis usadas.

Tal como os modelos anteriormente apresentados, as árvores de decisão também apresentam as suas vantagens e desvantagens. Em concordância com o mencionado por J. M. Pereira et al. (2007), as árvores de decisão caracterizam-se pela sua versatilidade e elevado índice de compreensão que permite identificar com mais facilidade os fatores mais importantes para classificar de forma mais assertiva as empresas em estudo.

Gepp et al. (2010) mencionam como vantagem das árvores de decisão o facto de as mesmas não imporem a necessidade de se usarem variáveis transformadas nem terem disposições passíveis de violar. Além disso, os mesmos acrescentam o facto de através das árvores de decisão os custos de classificação incorreta poderem ser incorporados ao processo de construção das mesmas, diminuindo os custos para a empresa.

No entanto, são encontradas algumas desvantagens a este tipo de modelo de previsão do fracasso empresarial nomeadamente exigirem probabilidades prévias de empresas fracassadas e não fracassadas como input na criação das mesmas, probabilidades essas

que são criadas de forma arbitrária e, portanto, constituem um fator de imprecisão. Além disso, acresce que as árvores de decisão apenas fornecem o significado relativo das variáveis, ao contrário dos modelos de previsão estatísticos (Gepp et al., 2010).

Já J. M. Pereira et al. (2007) mencionam como desvantagens o facto de o processo de criação das árvores de decisão ser demasiado lento, tal como se verificava nas redes neurais, e também o facto de nas árvores de decisão existir dificuldade de manuseamento da informação incompleta, bem como, o facto de poderem existir valores inesperados nos resultados deste método.

1.2.2.3 Máquinas Vetoriais de Suporte (SVM)

As máquinas vetoriais de suporte são exemplo de outro método baseado em inteligência artificial utilizado para prever o fracasso empresarial. De acordo com Alaka et al. (2018), as máquinas vetoriais de suporte utilizam um modelo linear criando um separador ótimo através de uma classe binária. Deste modo, segundo os mesmos autores, as variáveis mais próximas desse separador serão designadas por vetores de suporte e posteriormente utilizadas para definir o resultado, no caso empresa fracassada ou não fracassada.

De uma forma mais simplificada e seguindo a lógica de pensamento de Min & Lee (2005), as máquinas vetoriais de suporte transformam problemas complexos em problemas mais simples que possam ser passíveis de utilizar funções discriminantes.

Segundo Yeh et al. (2010), as máquinas vetoriais de suporte classificam-se, sobretudo, pela minimização do risco estrutural em relação às técnicas estatísticas, apresentando uma capacidade preditiva superior à dos métodos baseados em inteligência artificial anteriormente apresentados. Shetty et al. (2022) acrescentam como vantagem deste modelo de previsão do fracasso empresarial o facto de ser possível encontrar sempre um limite de decisão, ainda que exista sobreposição entre os dados utilizados.

Min & Lee (2005) referem ainda que as máquinas vetoriais de suporte se caracterizam pela sua elevada precisão e estabilidade em termos da previsão do fracasso empresarial. Além disso, trata-se de uma técnica de previsão bastante simples e que pode, eventualmente, ser utilizada em simultâneo com técnicas estatísticas, uma vez que

combina os pontos fortes dos métodos tradicionais, geralmente mais fáceis de analisar e mais relacionados com a teoria (Min & Lee, 2005).

A K-SVM – máquina vetorial de suporte baseada em Kernel – consiste num classificador não linear que utiliza diferentes funções de Kernel para efeitos de classificação (Shaw & Routray, 2017). Segundo os mesmos, a utilização de máquinas vetoriais de suporte baseadas em Kernel permite obter resultados mais fidedignos e realistas, apresentando uma melhoria da classificação quando comparados com outros modelos de previsão.

1.2.2.4 Algoritmo Genético (AG)

À semelhança dos modelos acima apresentados, o algoritmo genético é uma técnica de previsão do fracasso empresarial baseada em inteligência artificial desenvolvida por Holland (1975) citado por Gordini (2014) que consiste numa técnica de pesquisa estocástica que permite pesquisar uma quantidade de espaços mais complicados e maiores utilizando para tal efeito a ideia da genética natural e do princípio evolutivo.

Os algoritmos genéticos, segundo K.-S. Shin & Lee (2002), são mais adequados para problemas de otimização que tenham por base uma função objetivo sujeita a restrições rígidas e flexíveis. C.-H. Wu et al. (2007) acrescentam ainda que os algoritmos genéticos são adequados para a manipulação simultânea de modelos com diferentes resoluções na medida em que conseguem procurar espaços de solução não lineares sem precisarem de informação ou conhecimento prévio sobre as características do modelo. Enquanto que os modelos tradicionais direcionam para um ponto ótimo local, os algoritmos genéticos realizam a pesquisa em muitos pontos em simultâneo dando uma maior probabilidade para encontrar um ponto ótimo global (Gordini, 2014).

De acordo com K.-S. Shin & Lee (2002), os algoritmos genéticos realizam o processo de pesquisa através de quatro etapas – inicialização, seleção, cruzamento e mutuação. Inicialmente, na fase de inicialização, é realizada a seleção da população de estruturas genéticas designadas por cromossomas (no caso, base de dados) que são posteriormente distribuídas no espaço de soluções dando início à pesquisa. De seguida serão selecionados os melhores indivíduos através da função aptidão com vista a codificar numericamente o desempenho do cromossoma. Ainda nesta fase, os cromossomas que apresentem o maior

valor de aptidão serão selecionados e copiados para a geração seguinte levando a que os cromossomas com um melhor desempenho ocupem gradualmente uma porção cada vez maior da população. Na fase do cruzamento é realizado uma troca de segmentos entre os cromossomas correspondentes de uma cadeia, ou seja, pegar em várias soluções-mãe e produzi-las em soluções-filha. Por fim, a mutuação visa manter a diversidade genética de uma geração para a outra e consiste na introdução de alterações genéticas na população levando a que, eventualmente, se mude a solução obtida e podendo chegar a uma melhor solução (Gordini, 2014).

No contexto de previsão de falências, seguindo a linha de pensamento de K.-S. Shin & Lee (2002), para se obter a regra da falência é necessário encontrar pontos de corte para uma ou mais variáveis em análise, acima ou abaixo das quais uma empresa será classificada como “fracassada” ou “não fracassada”.

Estas técnicas de previsão do fracasso empresarial, ainda que não sejam tão comuns nem tão usadas como as apresentadas anteriormente, apresentam inúmeras vantagens quando comparados com as técnicas estatísticas nomeadamente o facto de serem mais exactos na previsão de dificuldades financeiras e apresentarem um melhor desempenho (C.-H. Wu et al., 2007).

Gordini (2014) refere como vantagens da utilização dos algoritmos genéticos para previsão do fracasso empresarial o facto de os AG serem de fácil adaptação, serem eficazes relativamente ao encontro do ponto ótimo global de uma dada função não linear e não ficam limitados a mínimos locais, assim como permitem obter regras fáceis de entender e de compreender. Outra das vantagens dos algoritmos genéticos para a previsão do fracasso empresarial deve-se ao facto de os mesmos terem a capacidade de aprender relações não lineares entre as variáveis de entrada (K.-S. Shin & Lee, 2002).

Contudo, este modelo apresenta também algumas limitações como o facto de, em concordância com K.-S. Shin & Lee (2002), apenas produzir previsões quando as regras estão ativadas enquanto que as redes neurais, por exemplo, realizam previsões em todos os casos.

Foram também realizados alguns estudos que comparam a capacidade preditiva dos algoritmos genéticos com outras técnicas nomeadamente com as máquinas vectoriais de suporte e a regressão linear como é o caso do estudo de Gordini (2014). Segundo o mesmo, foi possível verificar que os AG apresentam uma capacidade preditiva superior

quer sejam os modelos calculados tendo por base uma amostra agregada quer tenham por base uma amostra separada por grupos de dimensão.

No estudo realizado por C.-H. Wu et al. (2007) foi possível verificar que os modelos baseados em inteligência artificial, no caso o algoritmo genético, as máquinas vetoriais de suporte e as redes neurais, são mais exatos que os outros modelos estatísticos, concluindo que os modelos artificialmente inteligentes permitem classificar de forma mais correta as empresas fracassadas e não fracassadas.

Evidentemente, ao longo dos anos, foram desenvolvidas novas técnicas e modelos de previsão do fracasso empresarial, com o intuito de se aumentar a capacidade preditiva de cada método que podem ser adotados pelas empresas para preverem um possível insucesso da própria organização. Importa, contudo, ressaltar que apesar de toda essa evolução, atualmente a maioria das técnicas usadas continuam a ser as técnicas estatísticas por se tratarem de ferramentas relativamente simples e práticas de aplicar e se adaptar a qualquer estudo, bem como o facto de os dados utilizados pelas mesmas serem mais acessíveis por parte do utilizador do que as ferramentas de *data mining* usadas em técnicas de inteligência artificial (J. M. Pereira et al., 2007).

Convém mencionar também que as técnicas baseadas em inteligência artificial têm revelado uma tendência crescente de utilização podendo vir a ser uma alternativa às técnicas mais tradicionais no futuro.

CAPÍTULO II – O SETOR DOS TRANSPORTES

2 O Setor dos Transportes

O setor dos transportes, conforme mencionado mais em cima, é um setor crucial ao quotidiano de todos os indivíduos, permitindo transportar pessoas, bens e serviços entre diversos países assim como é um setor fundamental para as empresas europeias e para as cadeias de suprimentos globais (European Commission, 2023c).

A juntar a isto podemos acrescentar a importância deste setor para o comércio internacional, para o comércio nacional e para o desenvolvimento económico na medida em que permite transportar mercadorias de forma rápida, segura e económica (Eurostat, 2021).

Assume também um importante papel na economia do nosso país representando, seguindo a linha de pensamento de Murta (2010), um peso significativo do PIB nas contas nacionais. European Commission (2023a) menciona também que o setor dos transportes contribui para aproximadamente 5% do PIB da União Europeia e dá trabalho a mais de 10 milhões de pessoas em toda a Europa. No caso do nosso país e de acordo com o mencionado por Corselli-Nordblad et al. (2023), em 2021 estavam cerca de 136 000 pessoas empregadas no ramo dos transportes, a grande maioria, no caso 87% empregues no ramo dos transportes rodoviários de mercadorias. Trata-se, portanto, de um setor bastante produtivo e pré-cíclico na medida em que cresce com a economia e, simultaneamente, é um fator de crescimento para a economia (Murta, 2010).

Assistimos, por isso, diariamente a um crescimento rápido do transporte de mercadorias em toda a União Europeia que pode ser justificado pelo crescimento do comércio global e um conjunto de práticas económicas de que são exemplos as economias de escala e as entregas *just-in-time* (Eurostat, 2021).

2.1 Situação Atual do Setor

Assistimos, constantemente, a uma evolução deste setor como consequência da globalização das economias e da evolução dos mercados e das trocas entre países. Contudo, verifica-se atualmente uma crescente preocupação ambiental e, em concordância com European Commission (2023a), os transportes representam um peso significativo das emissões de gases com efeitos de estufa da União Europeia, no caso cerca de 25% do total de emissões, sendo um dos principais objetivos do *European Green*

Deal a redução em 90% dessas emissões. Deste modo, poderá o desenvolvimento do setor dos transportes estar comprometido a curto prazo, obrigando a que se adotem com mais frequência meios mais sustentáveis e amigos do ambiente ou, conforme o mencionado por Murta (2010), a mobilidade sustentável. Também Ala et al. (2021) referem que, dada a constante preocupação com a qualidade do ar, tem-se verificado um desenvolvimento de veículos mais sustentáveis e menos poluentes de que é exemplo a maior utilização do transporte ferroviário (European Commission, 2023b).

O setor dos transportes enfrenta também limitações associadas ao ruído, acidentes rodoviários, congestionamentos e, mais recentemente, a subida dos preços dos combustíveis associada ao conflito entre a Rússia e a Ucrânia (European Commission, 2023c).

Em suma, segundo a linha de pensamento de Eurostat (2021) pode o desenvolvimento e a evolução do setor dos transportes estar comprometida a curto prazo devido às constantes pressões sobre as infraestruturas de transportes nomeadamente ao nível de atrasos e congestionamentos, e também por força das questões de governação e da crescente preocupação ambiental.

2.2 Tipos de Transportes

Como é do conhecimento geral, o setor dos transportes pode ser dividido em três subsectores nomeadamente os transportes terrestres, os transportes marítimos e os transportes aéreos. Por sua vez, o transporte de mercadorias entre os vários países pode ser realizado por esses três meios que serão alvo de análise ao longo deste trabalho de investigação com especial destaque para o CAE 4941, 5020 e 5121.

De uma forma resumida, os transportes terrestres de mercadorias, englobam a deslocação física de mercadorias, em curta ou longa distância, realizada através de camiões, reboques, camiões-cisterna, bem como as operações de armazenagem e distribuição das próprias mercadorias. Dentro dos transportes terrestres podemos ainda dividir em transportes rodoviários de mercadorias, com um código da atividade económica de 4941, e os transportes ferroviários de mercadorias cujo código da atividade económica é 4920.

Segundo dados do Banco de Portugal, os transportes rodoviários de mercadorias são líderes no que diz respeito às trocas de mercadorias entre países, existindo em 2021, 8 252 empresas deste ramo no nosso país (Banco de Portugal, 2021). European Commission (2023a) assume também que os transportes terrestres representam o maior peso no transporte de mercadorias entre países, sendo mais de 50% da carga transportada por estrada na UE. No entanto, este meio de transporte é, simultaneamente, o mais poluente e que mais contribui para as emissões de gases com efeito de estufa.

O transporte rodoviário de mercadorias enfrenta, atualmente, alguns desafios associados à crescente preocupação ambiental levando a que se adotem meios mais sustentáveis e amigos do ambiente, como é o caso da mobilidade sustentável já referido anteriormente. Outro dos desafios associados a este modo de transporte prende-se com o facto de, em concordância com Corselli-Nordblad et al. (2023), ser muito associado a acidentes rodoviários, em muitas situações acidentes fatais, registando um rácio de 42 mortes por milhão de habitantes.

O transporte ferroviário de mercadorias, seguindo a linha de pensamento de Corselli-Nordblad et al. (2023), engloba o transporte físico de mercadorias através da utilização de um veículo ferroviário numa rede ferroviária.

Contudo, no caso do nosso país, o transporte ferroviário de mercadorias ainda não é muito usado tendo vindo a revelar um crescimento bastante positivo nos últimos anos podendo vir a ser uma aposta num futuro próximo dada a crescente preocupação ambiental. De acordo com European Commission (2023c), visa-se tornar o transporte terrestre de mercadorias mais sustentável e eficiente passando a otimizar as capacidades das estruturas ferroviárias disponíveis e apostar de forma crescente no transporte ferroviário.

Por sua vez, o transporte marítimo de mercadorias, cujo código da atividade económica é o 5020, consiste no transporte de mercadorias entre países através de navios de carga e em trajetos que são efetuados total ou parcialmente através do mar, bem como se incluem também o transporte de mercadorias para descarga no mar ou recuperados do mar e ainda o transbordo de um navio para o outro (Corselli-Nordblad et al., 2023). Segundo dados do Banco de Portugal, existiam 83 empresas de transporte marítimo de mercadorias em Portugal, em 2021 (Banco de Portugal, 2021).

O transporte marítimo de mercadorias é já um meio bastante utilizado para as trocas comerciais entre países e tem vindo a revelar uma tendência crescente ao longo dos anos, verificando-se, em 2021, cerca de 3,5 mil milhões de toneladas de mercadorias movimentadas pelos portos da União Europeia (Eurostat, 2022).

Neste caso e ao contrário do que se verificava nos transportes rodoviários de mercadorias, o transporte marítimo apresenta-se como bastante seguro no que diz respeito a acidentes registando um total de 24 mortes em 2021 (Corselli-Nordblad et al., 2023).

Por último, o transporte aéreo de mercadorias com um código de atividade económica 5121 consiste no transporte de mercadorias realizado através de meios aéreos, nomeadamente por aviões e, dado o rápido crescimento do transporte internacional, o transporte aéreo tem vindo a assumir uma crescente importância ao longo do tempo (Eurostat, 2021). Em concordância com o mencionado por Corselli-Nordblad et al. (2023), no nosso país o transporte aéreo de mercadorias não é muito significativo apresentando uma maior relevância em países como Alemanha, França e Países Baixos.

No caso de Portugal, de acordo com dados do Banco de Portugal, existiam 6 empresas do transporte aéreo de mercadorias em 2015, não existindo dados mais recentes sobre empresas deste ramo (Banco de Portugal, 2021).

3 Metodologia

Tendo este trabalho de investigação como objetivo principal a comparação do desempenho dos modelos de previsão do fracasso empresarial estatísticos e o desempenho dos modelos de previsão baseados em inteligência artificial, a metodologia adotada assenta em duas linhas: por um lado temos métodos baseados em estatística, no caso a regressão logística (LR) e, por outro lado, métodos baseados em inteligência artificial, nomeadamente as redes neurais (NN), as máquinas vetoriais de suporte (SVM) e as árvores de decisão (DT).

Deste modo, no estudo empírico desta investigação foi abordada a previsão do sucesso ou fracasso de uma empresa tendo por base um quadro de aprendizagem supervisionado e construindo um conjunto de dados relativos a um conjunto de indicadores financeiros das empresas e a respetiva informação correspondente a serem bem-sucedidas ou não. Tendo por base esse conjunto de indicadores financeiros foram construídos vários modelos para estudar o seu desempenho de previsão em novos dados não analisados nem vistos anteriormente de diferentes empresas.

Posteriormente, medimos a capacidade de previsão de cada modelo em termos das métricas de classificação habituais utilizadas pela comunidade de investigação em aprendizagem automática, nomeadamente a precisão, a recuperação, a pontuação *f-score* e a exatidão.

Importa aqui ressaltar que todos os dados foram retirados da base de dados Bureau Van Dijk – ORBIS e englobam dados contabilísticos de empresas num período compreendido entre 2014 e 2021.

3.1 Seleção das variáveis

Seguindo a linha de pensamento de J. M. Pereira et al. (2016) não existe uma base teórica que permita explicar o fracasso empresarial e simultaneamente sirva de orientação para o processo de seleção das variáveis explicativas, pelo que neste trabalho empírico se optou pela escolha de indicadores financeiros como sendo as nossas variáveis explicativas independentes e usadas para criar o modelo. Importa também mencionar que os rácios financeiros foram escolhidos tendo por base a sua maior utilização e o nível de significância obtido em alguns estudos anteriores.

A escolha de rácios financeiros como variáveis independentes para os modelos desenvolvidos deve-se ao facto de, seguindo a linha de pensamento de Lacher et al. (1995), os mesmos nos mostrarem a eficiência de uma empresa permitindo comparar as relações entre os valores contabilísticos que, no fundo, representam o resultado da empresa. Além disso, analisando a informação contabilística e financeira de uma empresa é possível tirar um “juízo” relativamente à saúde financeira e a perspectivas futuras de continuidade da mesma (Idrissi & Taouab, 2019). Altman (1968) acrescenta ainda que pelo facto de a previsão de eventuais dificuldades financeiras e de uma possível falência das empresas ser particularmente sensível à análise de rácios financeiros se deva adotar tais mecanismos para prever o sucesso ou o insucesso das organizações.

Existe um vasto conjunto de rácios financeiros que podem ser utilizados para análise do fracasso/sucesso empresarial. Segundo Lacher et al. (1995), os rácios financeiros podem ser agrupados em sete categorias, nomeadamente Rendibilidade do Investimento, Alavancagem Financeira, Rotação de Capital, Liquidez a Curto Prazo, Posição de Caixa, Rotação de Stocks e Rotação de Contas a Receber. Já Idrissi & Taouab (2019) referem que os rácios de estrutura financeira, a rendibilidade, a solvabilidade, a liquidez e o endividamento são os mais utilizados e que revelam um maior interesse por parte dos investigadores.

Neste caso, foi criado um conjunto de dados utilizando-se um conjunto de rácios de estrutura financeira, rendibilidade, solvabilidade, liquidez e endividamento. O conjunto de dados específico é, então, composto por 15 indicadores financeiros que representam as nossas variáveis independentes, nomeadamente ROE com base no resultado antes de impostos, ROA com base no resultado antes de impostos, ROE com base no resultado líquido, ROA com base no resultado líquido, margem de lucro, margem EBITDA, margem EBIT, fluxo de caixa/receitas operacionais, rotação do ativo líquido, rácio de liquidez, rácio de solvabilidade (baseado no ativo), endividamento, custos com pessoal/receitas operacionais.

3.2 Amostra e conjunto de dados

Conforme referido anteriormente, foi criado um conjunto de dados com 15 indicadores financeiros que caracterizam um total inicial de 5854 empresas subdivididas

por 4 países, nomeadamente Itália com 1580 empresas, Espanha com 1949, Portugal com 1216 e França com 1109 empresas.

O conjunto completo de dados inclui empresas fracassadas e empresas não fracassadas de modo a ser possível verificar a viabilidade dos modelos em análise. Assim, temos um total de 3869 empresas bem-sucedidas, o que corresponde a 66%, e 1985 empresas fracassadas correspondendo a 34% das empresas.

Em termos de variáveis independentes e em concordância com o supramencionado temos os indicadores financeiros e também o país de origem. Os indicadores financeiros e o país de origem foram considerados as nossas variáveis independentes e a variável indicadora de empresa fracassada ou empresa não fracassada foi considerada a nossa variável dependente. Posteriormente, optamos por descartar a utilização do país de origem como variável independente na medida em que utilizando-o como uma função indicadora permitiu obter resultados que não se alteravam significativamente com ou sem esse indicador.

Através da inclusão das variáveis independentes é possível obter uma função discriminante tanto para os modelos estatísticos como para os modelos baseados em inteligência artificial que nos permite verificar se as empresas em análise são fracassadas ou não e também fazer a comparação da capacidade preditiva dos modelos estatísticos e dos modelos baseados em inteligência artificial para as empresas em estudo.

3.3 Preparação dos dados

A exaustividade dos dados do conjunto de dados inicial era de cerca de 95,9% uma vez que para algumas das empresas não estavam presentes todos os indicadores escolhidos. Deste modo, para maximizar o número de pontos de dados que poderiam ser utilizados nos modelos e dado que a maioria dos dados em falta estava apenas associada a um dos indicadores financeiros – o rácio de solvência baseado no passivo – optamos por descartar este indicador financeiro para todas as empresas para realizar esta análise.

Assim, o conjunto final da amostra de dados inclui um total de 4866 empresas das quais 592 empresas são da Itália, 1949 são da Espanha, 1216 são de Portugal e as restantes 1109 empresas são da França. A amostra final é ainda dividida em empresas fracassadas, um total de 1985, e empresas não fracassadas no total de 2881 empresas.

Na fase seguinte, dividimos de forma aleatória o conjunto de dados num conjunto de treino e num conjunto de teste, 80% e 20% respetivamente, de forma estratificada e mantendo o rácio de empresas bem-sucedidas e falhadas em ambos os conjuntos. Utilizamos a mesma divisão de treino/teste para o treino e avaliação de todos os modelos que desenvolvemos.

Por fim, os dados foram normalizados através da utilização do método clássico da subtração da média e divisão pelo desvio padrão para cada variável independente, dentro do conjunto de treino.

CAPÍTULO IV – MODELOS E ANÁLISE DOS RESULTADOS

4 Modelos e Análise dos Resultados

Os indicadores financeiros, selecionados anteriormente, foram utilizados para construir vários modelos de aprendizagem automática para prever o sucesso ou fracasso das empresas e ser possível efetuar a comparação entre os modelos estatísticos e os modelos baseados em inteligência artificial.

4.1 Modelos Utilizados

Os métodos implementados foram: Máquinas Vetoriais de Suporte Linear (L-SVM), Máquinas Vetoriais de Suporte de Kernel (K-SVM), K-Nearest Neighbors, Regressão Logística (LR), árvores de decisão (DT), Random Forests (RF), Extremely Random Forests (ERF), AdaBoost e Redes Neurais (NN).

No caso dos modelos globais e à semelhança do verificado em da Silva et al. (2023), nos modelos globais, para cada um dos métodos tentámos várias combinações perfeitas de parâmetros para o algoritmo específico e, de seguida, utilizamos a melhor combinação para contruir os restantes modelos. No caso do modelo L-SVM, tentámos vários pesos e diferentes perdas e números de interações e relativamente às árvores de decisão tentámos diferentes profundidades e diferentes critérios para a medição de uma divisão em cada nó. Já para as redes neurais, tentámos diferentes números de camadas ocultas e diferentes números de neurónios em cada camada (da Silva et al., 2023).

4.2 Resultados e Análise

Conforme tem vindo a ser referido, os modelos foram treinados e experimentados em empresas de Itália, Espanha, Portugal e França. Os resultados dos modelos escolhidos para cada tipo de classificador estão resumidos na Tabela 1 abaixo apresentada.

Modelo	Accuracy	Precision C0	Recall C0	F-score C0	Precision C1	Recall C1	F-score C1
L-SVM	60.0%+-0.0%	50.9%+-0.0%	49.9%+-0.0%	50.4%+-0.0%	66.0%+-0.0%	66.9%+-0.0%	66.4%+-0.0%
K-SVM	71.6%+-0.0%	66.9%+-0.0%	59.9%+-0.0%	63.2%+-0.0%	74.3%+-0.0%	79.5%+-0.0%	76.8%+-0.0%
KNN	68.5%+-0.0%	63.7%+-0.0%	52.6%+-0.0%	57.7%+-0.0%	70.9%+-0.0%	79.4%+-0.0%	74.9%+-0.0%
LR	58.4%+-0.0%	49.1%+-0.0%	58.2%+-0.0%	53.3%+-0.0%	67.1%+-0.0%	58.6%+-0.0%	62.5%+-0.0%
DT	72.1%+-0.0%	69.0%+-0.1%	57.2%+-0.0%	62.5%+-0.0%	73.6%+-0.0%	82.3%+-0.1%	77.8%+-0.0%

RF	72.8%+-0.3%	71.1%+-0.6%	56.0%+-0.5%	62.7%+-0.4%	73.6%+-0.2%	84.3%+-0.5%	78.6%+-0.3%
ERF	73.4%+-0.6%	70.5%+-1.1%	59.6%+-0.8%	64.6%+-0.8%	74.9%+-0.4%	82.8%+-0.9%	78.6%+-0.6%
AdaBoost	71.8%+-0.0%	77.7%+-0.0%	43.1%+-0.0%	55.4%+-0.0%	70.0%+-0.0%	91.5%+-0.0%	79.3%+-0.0%
NN	71.9%+-0.7%	68.7%+-2.2%	57.5%+-2.3%	62.5%+-0.8%	73.7%+-0.6%	81.9%+-2.5%	77.6%+-0.9%

Tabela 1: Resultados dos modelos por cada tipo de classificador

Os resultados apresentados na Tabela 1 mostram-nos a média e o desvio-padrão das métricas para 10 execuções de cada método. Verificou-se, assim, que a maioria dos métodos é bastante estável no que diz respeito a diferentes execuções, com exceção das redes neurais que apresentam um desvio-padrão superior a 0,5% que pode ser justificado pela diferente inicialização aleatória dos parâmetros da rede nas diferentes execuções. Contudo, trata-se de um comportamento significativamente estável.

As métricas Precisão, Recuperação e Pontuação f foram calculadas para cada classe. Deste modo, dado o desequilíbrio no conjunto de dados, ou seja, o desequilíbrio entre empresas não fracassadas e empresas fracassadas, as métricas foram calculadas para empresas fracassadas (C0) e empresas não fracassadas (C1).

Em concordância com o referido no ponto anterior, as L-SVM foram desenvolvidos tendo por base pesos de classe equilibrados, com $C = 10$, 10 000 interações e uma perda de articulação. Os K-SVM foram treinados com um kernel RBF, com $C = 10$, pesos de classe equilibrados e um coeficiente de Kernel de escala. O K-NN foi desenvolvido tendo por base 5 vizinhos, algoritmo de árvore de bolas, pesos uniformes e a métrica de Manhattan. A regressão logística foi treinada em conjunto com o solucionados IBFGS, pesos de classe equilibrados e sem regularização. As DT foram desenvolvidas com uma profundidade de árvore 6 com todas as características, um critério de entropia e pesos de classe uniformes. O RF foi desenvolvido com 1 000 estimadores usando \log^2 do número de características para cada estimador, pesos de classe balanceados e critério do índice de Gini. Já o ERF foi treinado com 100 estimadores e utilizando todos os recursos, pesos de classe uniformes e o critério da entropia semelhante ao que se verificava nas árvores de decisão. O AdaBoost foi treinado com 10 estimadores e o algoritmo SAMME.R utilizando uma taxa de aprendizagem de 0,5. Por fim, as NN foram desenvolvidas com 3 camadas ocultas e uma saída softmax, 100 neurónios na primeira camada e 10 neurónios na segunda e terceira camadas, uma perda de entropia cruzada categórica esparsa sem

pesos de classe, o otimizador Adam a uma taxa de $10e^{-4}$, um tamanho de lote de 64 e 100 épocas de treino.

Como resultados da aplicação destes modelos nas 4866 empresas selecionadas temos que a maioria dos métodos alcançou níveis de precisão e exatidão na ordem dos 71% - 73%, tendo a melhor precisão sido obtida pelo método ERF.

Para prever o insucesso das empresas fracassadas, representada na tabela pela métrica C0, a melhor pontuação f foi também obtida pelo modelo ERF com o melhor equilíbrio entre a precisão e a recuperação. Contudo, neste caso a melhor precisão foi obtida com recurso ao AdaBoost e a melhor recuperação foi conseguida com recurso ao K-SVM.

Por sua vez, para as empresas não fracassadas, métrica C1 representada na tabela 1, a melhor pontuação f foi alcançada com recurso ao AdaBoost que, neste caso, registou também a maior recuperação. Já a melhor capacidade de precisão foi aqui conseguida com recurso ao ERF.

Assim, de forma resumida podemos afirmar que os modelos mais adequados para este estudo em causa são o ERF e o AdaBoost pois apresentam os melhores resultados, quer em termos de capacidade de precisão quer em termos de recuperação.

A previsão do fracasso empresarial é um tema que tem vindo cada vez mais e em sentido de concordância com aquilo que foi mencionado ao longo de todo este trabalho, a ser abordado com vista a se tornar numa ferramenta útil e que possa ser utilizado pelas empresas para preverem situações de falências (Shi & Li, 2019). Como tal, a pertinência e contributo deste estudo prende-se com o facto de compreender se os modelos baseados em inteligência artificial deverão ser uma aposta num futuro próximo de forma a colmatar lacunas dos modelos estatísticos e a obter com maior exatidão informações sobre uma possível falência das empresas.

A presente dissertação tinha como principal propósito comparar o desempenho dos modelos de previsão do fracasso empresarial estatísticos e os modelos de previsão do fracasso empresarial baseados em inteligência artificial no setor dos transportes num horizonte temporal compreendido entre 2014 e 2021, de modo a compreender qual o mais eficiente. Para tal, foi utilizado a base de dados Bureau Van Dijk – ORBIS com vista a retirar dados acerca de rácios financeiros, no caso um total de 15 indicadores financeiros que nos ajudaram a analisar a situação das empresas e a definir os modelos de previsão. Analisou-se, portanto, um conjunto de 4866 empresas divididas por quatro países: 592 empresas eram de Itália, 1949 eram de Espanha, 1216 eram de Portugal e 1109 empresas eram da França. O conjunto eram também constituído por empresas fracassadas, um total de 1985 empresas, e empresas não fracassadas, um total de 2881 empresas.

Inicialmente e servindo de ponto de partida para o desenvolvimento deste trabalho de investigação foi realizada uma breve revisão de literatura de modo a compreender e a analisar alguns modelos de previsão do fracasso empresarial propostos por alguns autores. São estudos que têm vindo a ser desenvolvidos ao longo dos anos dada a necessidade de, por um lado, colocar em prática os modelos existentes e, por outro lado, acompanhar a evolução empresarial. Conseguimos verificar que cada autor utiliza a seu próprio conceito relativamente a fracasso empresarial e adota o modelo de previsão do fracasso empresarial que for mais adequado para a situação em específico, não existindo por isso um conceito e um modelo único. Foi possível concluir que, apesar de toda a evolução a que temos assistido, os modelos tradicionais, tais como a análise discriminante múltipla, continuam a ser os mais utilizados verificando-se, no entanto, uma aposta crescente nos modelos baseados em inteligência artificial, nomeadamente as redes neurais e as máquinas vetoriais de suporte.

De seguida e incidindo este trabalho de investigação sobre o setor dos transportes, foi realizado um pequeno estudo sobre o setor por forma a compreender a situação atual do mesmo no nosso país e quais os principais desafios que enfrenta atualmente.

Com recurso ao *software* estatístico foram criados um conjunto de nove modelos: Máquinas Vetoriais de Suporte Linear (L-SVM), Máquinas Vetoriais de Suporte de Kernel (K-SVM), K-Nearest Neighbors, Regressão Logística (LR), árvores de decisão (DT), Random Forests (RF), Extremely Random Forests (ERF), AdaBoost e Redes Neurais (NN). Posteriormente, fizemos correr os modelos com base nos dados financeiros das empresas em análise de forma a analisar a capacidade preditiva dos mesmos.

Os resultados obtidos permitiram verificar que a maioria dos modelos utilizados apresenta níveis de precisão e exatidão elevados, compreendidos entre 71% e 73%, sendo o método ERF o que apresenta melhores resultados, quer em termos de capacidade preditiva quer em termos de recuperação mais preciso. Concluiu-se também que, quando comparados os modelos estatísticos de previsão do fracasso empresarial com os modelos baseados em inteligência artificial, os segundos apresentam uma capacidade preditiva superior, com especial destaque para o modelo AdaBoost e o ERF.

Relativamente aos indicadores financeiros, numa fase inicial consideramos também o país de origem como variável relevante para a classificação de uma empresa como fracassada ou não. No entanto foi possível concluir que o mesmo não é significativo para a esta classificação das empresas, uma vez que os resultados não se alteravam significativamente.

Quanto às limitações desta investigação podemos mencionar o facto de ser apenas direcionada para o setor dos transportes e a base de dados ser constituída por apenas quatro países, ainda que tenha sido uma base de dados suficiente e nos tenha permitido realizar esta análise e alcançar o objetivo proposto.

Para além disso, estamos a tratar de modelos de previsão pelo que os resultados obtidos quando utilizados para prever possíveis acontecimentos futuros, poderão não ser realistas e não se verificar, devendo ter sempre em atenção o fator da incerteza.

Podemos ainda mencionar como limitação deste trabalho de investigação o facto de os indicadores financeiros, utilizados como variáveis independentes, serem fornecidos pelas próprias empresas, correndo o risco de serem dados manipulados ou alterados.

Como proposta para trabalhos de investigação futuros que possam surgir nesta temática podemos sugerir a utilização de um conjunto relativamente mais reduzido de modelos, quer estatísticos quer baseados em inteligência artificial, por forma a ser possível realizar uma análise comparativa mais aprofundada nesta área.

Poderá também ser útil analisar a utilização de outros dados financeiros ou analisar os rácios financeiros de forma mais pormenorizada por forma a conseguir perceber se os mesmos foram alvo de manipulação contabilística antes de se passar para o teste e criação do modelo de previsão.

Outra proposta de investigação futura prende-se com o facto de se analisar se as empresas deverão, ou não, ser alertadas para eventuais situações de fracasso de modo a adotarem medidas e se conseguirem prevenir de tal situação.

Por fim, podemos então concluir que o objetivo a que nos propusemos no início desta investigação foi cumprido – os modelos de previsão do fracasso empresarial baseados em inteligência artificial apresentam uma capacidade preditiva superior aos modelos de previsão do fracasso empresarial estatísticos.

REFERÊNCIAS BIBLIOGRÁFICAS

- Ala, G., Colak, I., Di Filippo, G., Miceli, R., Romano, P., Silva, C., Valtchev, S., & Viola, F. (2021). Electric Mobility in Portugal: Current Situation and Forecasts for Fuel Cell Vehicles. *Energies* 2021, Vol. 14, Page 7945, 14(23), 7945. <https://doi.org/10.3390/EN14237945>
- Alaka, H. A., Oyedele, L. O., Owolabi, H. A., Kumar, V., Ajayi, S. O., Akinade, O. O., & Bilal, M. (2018). Systematic review of bankruptcy prediction models: Towards a framework for tool selection. *Expert Systems with Applications*, 94, 164–184. <https://doi.org/10.1016/J.ESWA.2017.10.040>
- Altman, E. I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *The Journal of Finance*, 23(4), 589–609. <http://www.jstor.org/about/terms.html>.
- Altman, E. I., & Hotchkiss, E. (2006). *Corporate Financial Distress and Bankruptcy - Predict and Avoid Bankruptcy, Analyze and Invest in Distressed Debt*.
- Altman, E. I., Marco, G., & Varetto, F. (1994). Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking & Finance*, 18(3), 505–529. [https://doi.org/10.1016/0378-4266\(94\)90007-8](https://doi.org/10.1016/0378-4266(94)90007-8)
- Amankwah-Amoah, J., Adomako, S., & Berko, D. O. (2022). Once bitten, twice shy? The relationship between business failure experience and entrepreneurial collaboration. *Journal of Business Research*, 139, 983–992. <https://doi.org/10.1016/J.JBUSRES.2021.10.044>
- Amankwah-Amoah, J., & Wang, X. (2019). Business Failures around the World: Emerging Trends and New Research Agenda. *Journal of Business Research*, 98, 367–369. <https://doi.org/10.1016/J.JBUSRES.2019.02.064>
- Banco de Portugal. (2021). *Quadros do Setor*. <https://www.bportugal.pt/QS/qsweb/Dashboards>
- Beaver, W. H. (1966). Financial Ratios As Predictors of Failure. *Journal of Accounting Research*, 4, 71. <https://doi.org/10.2307/2490171>

- Beaver, W. H., Correia, M., & McNichols, M. F. (2010). Financial statement analysis and the prediction of financial distress. *Foundations and Trends in Accounting*, 5(2), 99–173. <https://doi.org/10.1561/1400000018>
- Borchert, P., Coussement, K., De Caigny, A., & De Weerd, J. (2023). Extending business failure prediction models with textual website content using deep learning. *European Journal of Operational Research*, 306(1), 348–357. <https://doi.org/10.1016/j.ejor.2022.06.060>
- Corselli-Nordblad, L., Jere, N., Ford-Alexandraki, E., Xenellis, G., & Statistical Office of the European Communities. (2023). *Key figures on European transport : 2022 edition*. (N. Jere, L. Corselli-Nordblad, E. Ford-Alexandraki, & G. Xenellis, Eds.). Publications Office of the European Union. <https://doi.org/10.2785/322262>
- da Silva, A. F., Brito, J. H., & Pereira, J. M. (2023). Using Machine Learning to Predict Business Failure in Iberian Hospitality Sector. In *Advances in Tourism, Technology and Systems* (Vol. 2, pp. 313–322).
- Deakin, E. B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research*, 10(1), 167. <https://doi.org/10.2307/2490225>
- Dong, J. R., & Chen, J. (2009). An evolutionary based wavelet network for business failure prediction. *Proceedings of the 2009 WRI Global Congress on Intelligent Systems, GCIS 2009*, 3, 274–278. <https://doi.org/10.1109/GCIS.2009.226>
- European Commission. (2023a). *Green Deal: Greening freight for more economic gain with less environmental impact**.
- European Commission. (2023b). *Sustainable transport*. Sustainable Transport. https://transport.ec.europa.eu/transport-themes/sustainable-transport_en
- European Commission. (2023c). *Transport and the Green Deal*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal/transport-and-green-deal_en
- Eurostat. (2021). *Freight transport statistics - Statistics Explained*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Archive:Freight_transport_statistics#Maritime_freight

- Eurostat. (2022, October). *Maritime freight and vessels statistics - Statistics Explained*.
https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Maritime_freight_and_vessels_statistics
- Gámez, M. A. F., Gil, A. C., & Ruiz, A. J. C. (2016). Applying a probabilistic neural network to hotel bankruptcy prediction. *Tourism & Management Studies*, 12(1), 40–52. <https://doi.org/10.18089/tms.2016.12104>
- Gepp, A., Kumar, K., & Bhattacharya, S. (2010). Business failure prediction using decision trees. *Journal of Forecasting*, 29(6), 536–555. <https://doi.org/10.1002/FOR.1153>
- Gordini, N. (2014). A genetic algorithm approach for SMEs bankruptcy prediction: Empirical evidence from Italy. *Expert Systems with Applications*, 41(14), 6433–6445. <https://doi.org/10.1016/j.eswa.2014.04.026>
- Hauser, R. P., & Booth, D. (2022). Predicting Bankruptcy with Robust Logistic Regression. *Journal of Data Science*, 9(4), 565–584. [https://doi.org/10.6339/JDS.201110_09\(4\).0006](https://doi.org/10.6339/JDS.201110_09(4).0006)
- Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, 28(4), 331–343. <https://doi.org/10.1002/SMJ.582>
- Idrissi, K., & Taouab, O. (2019). Towards a Financial-based View of Business Failure. *Revue Internationale Des Sciences de Gestion*, 2, 154–172. www.revue-isg.com
- Korol, T., & Spyridou, A. (2020). Examining Ownership Equity as a Psychological Factor on Tourism Business Failure Forecasting. *Frontiers in Psychology*, 10, 3048. <https://doi.org/10.3389/FPSYG.2019.03048/BIBTEX>
- Lacher, R. C., Coats, P. K., Sharma, S. C., & Fant, L. F. (1995). A neural network for classifying the financial health of a firm. *European Journal of Operational Research*, 85(1), 53–65. [https://doi.org/10.1016/0377-2217\(93\)E0274-2](https://doi.org/10.1016/0377-2217(93)E0274-2)
- Lennox, C. (1999). Identifying failing companies: A reevaluation of the logit, probit and DA approaches. *Journal of Economics and Business*, 51(4), 347–364. [https://doi.org/10.1016/S0148-6195\(99\)00009-0](https://doi.org/10.1016/S0148-6195(99)00009-0)

- Maricica, M., & Georgeta, V. (2012). Business Failure Risk Analysis using Financial Ratios. *Procedia - Social and Behavioral Sciences*, 62, 728–732. <https://doi.org/10.1016/J.SBSPRO.2012.09.123>
- Min, J. H., & Lee, Y. C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4), 603–614. <https://doi.org/10.1016/J.ESWA.2004.12.008>
- Murta, D. F. V. (2010). *Quilómetros, euros e pouca terra: manual de economia dos transportes*. Imprensa da Universidade de Coimbra. <https://doi.org/10.14195/978-989-26-0492-3>
- Neves, J. C., & Vieira, A. (2006). Improving bankruptcy prediction with Hidden Layer Learning Vector Quantization. *European Accounting Review*, 15(2), 253–271. <https://doi.org/10.1080/09638180600555016>
- Ohlson, J. A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 18(1), 109–131.
- Ooghe, H., & Spaenjers, C. (2009). A note on performance measures for business failure prediction models. *Applied Economics Letters*, 17(1), 67–70. <https://doi.org/10.1080/13504850701719769>
- Pereira, J., Basto, M., & Ferreira-da-Silva, A. (2014). Comparative Analysis between Statistical and Artificial Intelligence Models in Business Failure Prediction. *Journal of Management and Sustainability*, 4(1). <https://doi.org/10.5539/jms.v4n1p114>
- Pereira, J. M., Ángel, M., Domínguez, C., Luís, J., & Ocejo, S. (2012). LA PREDICCIÓN DEL FRACASO EMPRESARIAL. PROPUESTA DE UN MODELO SECUENCIAL BASADO EN EL ANÁLISIS DE SUPERVIVENCIA. XV *Encuentro de AECA*.
- Pereira, J. M., Basto, M., Díaz Gómez, F., Titular, P., & Albuquerque, E. B. (2010). *LOS MODELOS DE PREDICCIÓN DEL FRACASO EMPRESARIAL. PROPUESTA DE UN RANKING*.
- Pereira, J. M., Basto, M., & Ferreira Das Silva, A. (2017). *Comparing logit model with discriminant analysis for predicting bankruptcy in Portuguese hospitality sector*.

- Pereira, J. M., Crespo Domínguez, M. Á., & Ocejo, J. L. S. (2007). Modelos de Previsão do Fracasso Empresarial: Aspectos a considerar. *Revista de Estudos Politécnicos Polytechnical Studies Review*, 7, 111–148.
- Pereira, J. M., Ferreira da Silva, A., Basto, M., & José Palma Lampreia dos Santos, M. (2016). *A previsão do fracasso empresarial utilizando a análise discriminante e o logit no setor hoteleiro português.*
- Shaw, L., & Routray, A. (2017). A critical comparison between SVM and k-SVM in the classification of Kriya Yoga meditation state-allied EEG. *WIECON-ECE 2016 - 2016 IEEE International WIE Conference on Electrical and Computer Engineering*, 134–138. <https://doi.org/10.1109/WIECON-ECE.2016.8009103>
- Shetty, S., Musa, M., & Brédart, X. (2022). Bankruptcy Prediction Using Machine Learning Techniques. *Journal of Risk and Financial Management*, 15(1). <https://doi.org/10.3390/jrfm15010035>
- Shi, Y., & Li, X. (2019). An overview of bankruptcy prediction models for corporate firms: A systematic literature review. *Intangible Capital*, 15(2), 114–127. <https://doi.org/10.3926/ic.1354>
- Shin, K. S., & Lee, Y. J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321–328. [https://doi.org/10.1016/S0957-4174\(02\)00051-9](https://doi.org/10.1016/S0957-4174(02)00051-9)
- Shin, K.-S., & Lee, Y.-J. (2002). A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications*, 23(3), 321–328. [https://doi.org/10.1016/S0957-4174\(02\)00051-9](https://doi.org/10.1016/S0957-4174(02)00051-9)
- Tam, K. Y., & Kiang, M. Y. (1992). Managerial Applications of Neural Networks: The Case of Bank Failure Predictions. <https://doi.org/10.1287/Mnsc.38.7.926>, 38(7), 926–947. <https://doi.org/10.1287/MNSC.38.7.926>
- Walsh, G. S., & Cunningham, J. A. (2016). Business Failure and Entrepreneurship: Emergence, Evolution and Future Research. *Foundations and Trends® in Entrepreneurship*, 12(3), 163–285. <https://doi.org/10.1561/03000000063>
- Wu, C.-H., Tzeng, G.-H., Goo, Y.-J., & Fang, W.-C. (2007). A real-valued genetic algorithm to optimize the parameters of support vector machine for predicting

bankruptcy. *Expert Systems with Applications*, 32(2), 397–408.
<https://doi.org/10.1016/j.eswa.2005.12.008>

Wu, W. W. (2010). Beyond business failure prediction. *Expert Systems with Applications*, 37(3), 2371–2376. <https://doi.org/10.1016/J.ESWA.2009.07.056>

Yeh, C. C., Chi, D. J., & Hsu, M. F. (2010). A hybrid approach of DEA, rough set and support vector machines for business failure prediction. *Expert Systems with Applications*, 37(2), 1535–1541. <https://doi.org/10.1016/J.ESWA.2009.06.088>

Zmijewski, M. E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research*, 22, 59.
<https://doi.org/10.2307/2490859>