

A IMPORTÂNCIA DA MATURIDADE DA DÍVIDA NA
TOMADA DE DECISÃO NAS PME EUROPEIAS –
EVIDÊNCIA DE PORTUGAL, ESPANHA, FRANÇA, ITÁLIA,
ALEMANHA E GRÉCIA

Tiago Henrique Filipe Monteiro

Dissertação submetida como requisito parcial para obtenção do grau de
Mestre em Finanças

Orientador:

Prof. Doutor Raul Laureano, Prof. Auxiliar, ISCTE Business School, Departamento de
Métodos Quantitativos para Gestão e Economia

Coorientador:

Prof. Doutor João Guerreiro, Prof. Auxiliar Convidado, ISCTE Business School,
Departamento de Marketing, Operações e Gestão Geral

Setembro de 2015

Epígrafe

*“Aprender é a única coisa de que a mente nunca se cansa,
nunca tem medo e nunca se arrepende.”*

Leonardo Da Vinci

*“Para vencer - material ou imaterialmente –
três coisas definíveis são precisas:
saber trabalhar, aproveitar oportunidades, e criar relações”*

Fernando Pessoa

Agradecimentos

Um profundo obrigado.

Em primeiro lugar aos orientadores, o Prof. Doutor Raul Laureano e o Prof. Doutor João Guerreiro pela dedicação e empenho que disponibilizaram desde o primeiro momento, pela partilha do conhecimento, profissionalismo e entusiasmo que foram essenciais nas horas mais difíceis.

A todos os professores do Mestrado em Finanças que partilharam todo o seu conhecimento, conhecimento esse, fundamental para a elaboração da presente tese.

Aos meus pais, pelo apoio e carinho que me deram numa das fases mais importantes da minha vida, que se revelou essencial para a elaboração da presente dissertação. Agradeço-lhes a sua compreensão pelas horas de isolamento da minha parte e terem sofrido as consequências das minhas limitações de disponibilidade e pela constante motivação com que pude contar ao longo destes meses.

Aos amigos pela força que sempre transmitiram e pela total compreensão da minha ausência.

A todos os meus colegas do ISCTE e INDEG com quem partilhei momentos memoráveis e que fizeram que fosse possível efetuar este trabalho.

O meu profundo e sentido agradecimento a todas as pessoas que contribuíram para a concretização desta tese.

Resumo

A estrutura de capitais e a maturidade da dívida das empresas tem constituído, nas últimas décadas, um dos temas de maior interesse na literatura financeira. Desde o estudo de Modigliani & Miller (1958), vários estudos e teorias têm sido desenvolvidas nesta área, inicialmente aplicadas nas grandes empresas, e entretanto também aplicadas ao universo das PME.

O principal objetivo desta investigação é analisar quais os fatores que mais contribuem para a tomada de decisão por parte dos gestores relativamente à maturidade da dívida nas PME Europeias.

A análise baseou-se na Metodologia CRISP-DM, recorrendo-se na fase da modelação a diversas técnicas de regressão, nomeadamente Regressão Linear Múltipla, Árvores de Decisão e Redes Neurais Artificiais. A amostra é composta por 86.158 empresas de Portugal, Espanha, Itália, Alemanha, França e Grécia, com dados referentes ao ano 2013, extraídos da base de dados Amadeus.

Esta investigação apresenta-se como uma das primeiras a fazer a comparação de diferentes PME de países da União Europeia bem como a utilizar técnicas de DM para a explicação da maturidade da dívida. As Árvores de Decisão e as Redes Neurais Artificiais apresentam resultados consideravelmente superiores aos demonstrados na Regressão Linear Múltipla na previsão da maturidade da dívida de curto e, também, de longo prazo.

As conclusões apontam para que o risco, a estrutura dos ativos e a alavancagem sejam os fatores que melhor explicam a tomada de decisão por parte dos gestores relativamente à maturidade da dívida. Os resultados demonstram ainda que a localização, o setor de atividade e tipo de empresa não são importantes para a explicação da maturidade da dívida.

Palavras-Chave: Árvores de Decisão, *Data Mining*, Estrutura de capitais, Maturidade da Dívida; Pequenas e Médias Empresas (PME), Redes Neurais Artificiais e Regressão.

Classificação JEL: G01, G30, G32; G39

Abstract

Since the past decade the Capital Structure and more specifically the debt maturity has become, one of the topics of greatest interest in financial literature. After the Modigliani & Miller (1958) study publication, several studies and theories have been developed in this area. Those studies were initially applied in large companies and only applied to the universe of SME's afterwards.

The purpose of this research is to analyze which factors mostly contribute to the decision making by managers in relation of debt maturity according in European SME's.

The analysis was based on the CRISP-DM Methodology, with the modeling approach of using different regression techniques, including Multiple Linear Regression; Decision Trees and Artificial Neural Networks. The sample is composed of 86.158 companies from Portugal, Spain, Italy, Germany, France and Greece, in 2013, and the data was extracted from Amadeus database.

This research appears as one of the firsts studying the debt maturity using the comparison between different SME's in the European Union countries and using data mining techniques. The Decision Trees and Artificial Neural Networks show considerably better results than those shown in Multiple Linear Regression.

The conclusions indicate that the factors that better explain the managers' decision making processes regarding the debt to maturity are the risk, the structure of assets and the leverage. The results also show that the location, the business sector and the type of company are not important to explain the debt maturity.

Keywords: Decision Trees, Data Mining, Capital Structure, Debt Maturity, Small and Medium Enterprises (SME's), Artificial Neural Network and Regression.

JEL Classification: G01, G30, G32; G39

Abreviaturas

AD – Árvores de Decisão

AID – Automatic Interaction Detection

BI – Business Intelligence

CART – Classification and Regression Trees

CHAID – Chi-squared Automatic Interaction Detection

CMP – Custo Médio Ponderado

CRISP –DM – Cross Industry Standard Process for Data Mining

DCBD – Descoberta de Conhecimento em Bases de Dados

DCP – Dívida de Curto Prazo

DLP – Dívida de Longo Prazo

DM – Data Mining

DW – Data Warehouses

EAM – Erro Absoluto Médio

EQM – Erro Quadrático Médio

EUA – Estados Unidos da América

LSD – Least Squared Deviation

M&M – Modigliani & Miller

MAE – Mean Absolute Error

MLP – Multi-Layer Perceptron

MMQ – Método dos Mínimos Quadrados

MSE – Mean Squared Error

NYSE – New York Stock Exchange

OC – Oportunidades de Crescimento

OLAP – On-line Analytical Processing

PME – Pequenas e Médias Empresas

REC – Regression Error Characteristics

RMSE – Root Mean Squared Error

RNA – Redes Neurais Artificiais

RLM – Regressão Linear Múltipla

RQEQM – Raiz Quadrada do Erro Quadrático Médio

SEMMA – Sample, Explore, Modify, Model, Assessment

TI – Tecnologias da informação

TIC – Tecnologias da Informação e Comunicação

UE – União Europeia

VAL – Valor Atual Líquido

Índice

Capítulo 1 – Introdução	1
1.1. Âmbito	1
1.2. Objetivos	2
1.3. Abordagem Metodológica	3
1.4. Estrutura	4
Capítulo 2 – Estado da Arte	5
2.1. Referencial Teórico sobre a Estrutura de Capital	5
2.2. Caracterização das PME	7
2.3. Teorias da estrutura de capitais	8
2.3.1. Abordagem Tradicional	8
2.3.2. Irrelevância da Estrutura de Capitais	9
2.3.3. Teoria do Trade-off	10
2.3.4. Teoria da Agência	11
2.3.5. Teoria do Pecking Order	12
2.4. Fatores determinantes da maturidade da dívida	13
2.4.1. Dimensão das Empresas	13
2.4.2. Estrutura dos Ativos	14
2.4.3. Oportunidades de Crescimento	15
2.4.4. Rendibilidade	16
2.4.5. Risco	17
2.4.6. Impostos	17
2.4.7. Alavancagem	18
2.4.8. Cash-Flow	18
2.4.9. Setor de Atividade	19
2.5. Estudos Relevantes	19
2.6. Hipóteses de Investigação	21
Capítulo 3 – Descoberta de Conhecimento em Bases de Dados	25
3.1. <i>Business Intelligence</i>	25
3.2. Processo de Descoberta de Conhecimento em Bases de Dados	26
3.2.1. Fases do Processo de DCBD	27
3.3. Metodologias do <i>Data Mining</i>	28
3.3.1. Metodologia CRISP-DM	28
3.4. <i>Data Mining</i>	29
3.4.1. Paradigmas e tarefas de Data Mining	30
Capítulo 4 – Metodologia	33
4.1. Compreensão do problema	33
4.2. Compreensão e Preparação dos dados	34
4.3. Modelação	43
4.3.1. Regressão Linear Múltipla	44
4.3.2. Árvores de Decisão	50

4.3.3. Redes Neurais Artificiais.....	57
Capítulo 5 – Avaliação e Resultados.....	67
5.1. Avaliação.....	67
5.2. Implementação	69
5.3. Discussão e Resultados	70
Capítulo 6 – Conclusão	73
6.1. Principais Conclusões	73
6.2. Importância e Contributo do Estudo	73
6.3. Principais Limitações e Investigações Futuras	74
Referências Bibliográficas	77
Anexos.....	85

Índice de Tabelas

Tabela 1 – Rácios da Estrutura de Capitais	6
Tabela 2 – Enquadramento das PME	7
Tabela 3 – Principais estudos empíricos e as relações encontradas	20
Tabela 4 – Técnicas de DM e o tipo de problemas que solucionam	32
Tabela 5 – Principais estudos da maturidade da dívida e dimensão da amostra	35
Tabela 6 – Indicadores de Maturidade da Dívida.....	36
Tabela 7 – Fatores Determinantes da maturidade da dívida.....	38
Tabela 8 – Estatística Descritiva dos Indicadores da Maturidade	40
Tabela 9 – Estatística Descritiva dos Fatores Determinantes da Maturidade da Dívida	41
Tabela 10 – Correlação de Pearson entre os indicadores da maturidade e os fatores determinantes	42
Tabela 11 – Média dos Indicadores da Maturidade de acordo com a Dimensão das Empresas	43
Tabela 12 – Média dos Indicadores da Maturidade de acordo com o setor de atividade	43
Tabela 13 – Resultados da Regressão – Modelo Global	46
Tabela 14 – Impacto dos fatores determinantes na maturidade da dívida.....	50
Tabela 15 – Importância das variáveis	65
Tabela 16 – Resultados das Redes Neurais	66
Tabela 17 – Métricas Obtidas para a Maturidade de Curto Prazo.....	69

Índice de Figuras

Figura 1 – Processo de DCBD.....	27
Figura 2 – Metodologia CRIP – DM.....	28
Figura 3 – Paradigmas do Data Mining.....	30
Figura 4 – Exemplo de uma AD.....	51
Figura 5 – Exemplo da Ortogonalidade.....	54
Figura 6 – Redes Multicamadas.....	59
Figura 7 – Constituição do Neurónio.....	60
Figura 8 – Fase de Propagação.....	62
Figura 9 – Fase da <i>Backpropagation</i>	62

Capítulo 1 – Introdução

O principal objetivo deste capítulo introdutório é providenciar uma panorâmica global do trabalho que vai ser desenvolvido nos capítulos seguintes. No âmbito é feita uma breve introdução à importância das pequenas e médias empresas (PME) bem como à problemática da maturidade da dívida que é o propósito da presente investigação. De seguida são definidos os principais objetivos bem como a metodologia aplicada para os atingir. No final do capítulo é apresentada a estrutura da tese.

1.1. Âmbito

A globalização financeira e a volatilidade dos mercados de capitais fazem com que as empresas tenham de saber lidar com a mudança, sendo impreterível que as empresas consigam atingir a eficiência, não só em termos operacionais, mas também em termos financeiros. As decisões tomadas pelos gestores a nível operacional e principalmente a nível financeiro apresentam-se como críticas para o sucesso e sobrevivência das empresas.

As empresas de menor dimensão apresentam maiores dificuldades no acesso a financiamento do que as empresas de maior dimensão. As PME recorrem maioritariamente ao financiamento interno bem como a instituições de crédito informais, fundamentalmente instituições bancárias (Demirguç-Kunt & Maksimovic, 1998). Dada a importância que as PME apresentam na economia mundial, é fundamental estudar-se como é que estas entidades financiam os seus ativos e investigar quais os fatores que mais afetam as tomadas de decisão dos gestores no que diz respeito à criação de valor, retorno para os proprietários e definição da maturidade da dívida. Assim, este estudo permitirá atualizar a literatura existente com a realidade das PME europeias.

A dicotomia entre o capital próprio e a dívida começou a ser estudada na década de 50 do século passado a partir dos artigos seminais de Modigliani e Miller (1958, 1963). Foi a partir dos pressupostos propostos por Modigliani & Miller (1958) que apareceram diversas teorias, que foram sendo testadas junto das grandes empresas, nomeadamente, a Teoria da Agência, a Teoria do Trade-off e a Teoria do *Pecking Order*. Mais recentemente, as investigações sobre os fatores que afetam a maturidade da dívida começaram a incluir as PME. Os fatores mais importantes que explicam a maturidade da dívida sugeridos na literatura são: a estrutura dos

ativos, as oportunidades de crescimento; risco; alavancagem; impostos; rentabilidade; dimensão e os *cash flow*.

A escassez de literatura relacionada com a problemática da maturidade da dívida nas PME e a sua comparação com outros países consubstancia-se num dos principais fatores de motivação para o desenvolvimento deste trabalho académico, e, dentro do panorama atual, é um dos primeiros estudos, se não o primeiro, onde são aplicadas Árvores de Decisão (AD) e Redes Neurais Artificiais (RNA) no sentido de identificar que fatores são mais valorizados pelos gestores aquando da tomada de decisão.

1.2. Objetivos

No contexto da importância do estudo da maturidade da dívida, com destaque para a inclusão de PME de Portugal, Espanha, Grécia, Itália, França e Alemanha são delineados os seguintes objetivos: em primeiro lugar pretende-se extrair conhecimento útil para a complementar a literatura já existente identificando que fatores mais contribuem para a tomada de decisões por parte dos gestores relativamente à maturidade da dívida e determinar se os fatores que influenciam a maturidade da dívida são semelhantes para as decisões tomadas para o curto prazo e para o longo prazo.

No segundo objetivo pretende-se estudar se a maturidade da dívida varia de país para país, de setor para setor de atividade e por tipo de empresa.

Para estudar os objetivos anteriores vão ser aplicadas técnicas de *data mining* (DM), para comparar com os resultados obtidos com modelos de regressão linear múltipla, técnica utilizada na maioria dos estudos existentes, e perceber se é a mais adequada para estudar o fenómeno da maturidade da dívida nas PME. As técnicas de *DM* são ainda importantes no sentido de se perceber se é possível construir modelos que apresentem melhores resultados com um menor número de variáveis e que ao mesmo tempo auxiliem os gestores na tomada de decisões.

A identificação dos fatores que explicam a maturidade da dívida são importantes no contexto das PME dado que as decisões tomadas afetam a continuidade/estabilidade das empresas. Com a identificação dos fatores pretende-se auxiliar os gestores na tomada de decisão relativamente às questões relacionadas com o financiamento das PME.

1.3. Abordagem Metodológica

Qualquer trabalho académico requer inicialmente um levantamento bibliográfico sobre a problemática a estudar. De modo a concretizarem-se os objetivos delineados no ponto anterior definiu-se a metodologia CRISP-DM, muito útil em projetos de DM. Esta metodologia é bastante intuitiva e proporciona uma visão integrada e delimitada das seis etapas que um projeto de DM deverá adotar.

Seguindo a metodologia do CRISP-DM, iniciar-se-á o processo pela compreensão do problema, que tem subjacente a definição do problema bem como os objetivos a atingir. A fase seguinte compreende a compreensão dos dados, sendo nesta etapa que se faz uma análise à qualidade dos dados. A terceira etapa, preparação dos dados, é uma etapa essencial para o sucesso dos modelos de DM. É nesta etapa que se selecionam, transformam e limpam os dados. A quarta etapa, modelação, é onde se utilizam os diferentes algoritmos de DM – Regressão Linear Múltipla (RLM); as Árvores de Decisão (AD), e as Redes Neurais Artificiais (RNA). A etapa seguinte é composta pela avaliação dos resultados obtidos. Nesta fase comparam-se os resultados com o conhecimento já existente sendo ainda feita uma revisão de todos os passos efetuados para a construção dos modelos, confirmando se os modelos se enquadram nos objetivos delineados na fase da compreensão do problema. Por fim, a etapa da implementação, pode ser a construção de um simples relatório, ou algo mais complexo como a criação de técnicas de apoio à decisão que permita corroborar as decisões tomadas com base no conhecimento adquirido no processo.

Este estudo tem por base a seleção de uma amostra aleatória de 86.158 PME de Portugal, Espanha, Itália, Grécia, França e Alemanha, para o ano de 2013 na plataforma Amadeus (Amadeus, 2015). Esta plataforma contém informação financeira detalhada, principalmente ao nível das demonstrações financeiras e rácios financeiros bem como alguns dados qualitativos de cada empresa. Estes países foram selecionados porque pretende-se avaliar se a maturidade da dívida das empresas dos países mais afetados pela crise (Portugal, Espanha, Grécia e Itália) é idêntica à dos países menos afetados (França e Alemanha) e quais os fatores que mais influenciaram a tomada de decisão por parte dos gestores no ano de 2013.

1.4. Estrutura

A tese está organizada em seis capítulos. Após esta introdução, no capítulo 2 é feita uma breve apresentação das principais teorias relacionadas com a estrutura de capitais bem como dos fatores determinantes da maturidade da dívida. O capítulo 3 enquadra a temática da Descoberta de Conhecimento em Bases de Dados e de uma das suas principais fases, o DM, assim como da metodologia CRISP-DM. No capítulo 4 e de acordo com a metodologia CRISP-DM é abordada a parte empírica, fase onde são expostos os modelos de *DM*. No capítulo 5 é feita a comparação dos modelos bem como a apresentação dos principais resultados. Por fim, no capítulo 6 são apresentadas as principais conclusões do estudo empírico e quais as principais limitações inerentes ao estudo.

Capítulo 2 – Estado da Arte

Neste capítulo apresentam-se as principais teorias relacionadas com a problemática da maturidade da dívida. O capítulo é composto por uma breve explicação de cada uma das teorias complementando-as com os estudos empíricos mais relevantes. Em suma, neste capítulo realiza-se uma breve resenha sobre a evolução do estudo da maturidade da dívida e identifica-se o *background* que está na base do estudo empírico.

2.1. Referencial Teórico sobre a Estrutura de Capital

As empresas apresentam duas formas de financiamento: capital próprio e/ou capital alheio. Os capitais próprios e os capitais alheios apresentam várias características diferenciadoras, como sejam o tipo e estatuto fiscal de rendimento proporcionado bem como o controlo que os detentores têm sobre a entidade e os direitos em caso de falência da mesma (Mota, Barroso, Nunes, & Ferreira, 2012). O rendimento proporcionado pelo capital próprio são os dividendos, no entanto não é obrigatória a sua amortização, salvo algumas exceções previamente definidas no contrato das sociedades. Por outro lado, a opção por capital alheio tem subjacente um plano de serviço de dívida previamente fixado, plano este que apresenta maturidade e taxa de juro definidas (de Miguel & Pindado, 2001).

O capital alheio leva ainda ao acréscimo do retorno esperado pelos proprietários, no entanto também aumenta o risco e a probabilidade de insolvência/falência das empresas (Matias & Esperança, 2009). Esta situação não se verifica com o capital próprio na medida em que se a empresa não conseguir corresponder às expectativas dos proprietários não se encontra numa situação de falência.

As origens de capital podem ser divididas consoante a sua maturidade, ou seja, em curto e longo prazo. A maturidade de curto prazo referem-se a financiamentos para fazer face à atividade ordinária da empresa sendo a sua definição relativamente unânime na literatura, ou seja, é definida como a dívida com prazo inferior a um ano, onde se incluem principalmente o crédito de clientes, cartões de crédito e débito, contas caucionadas, cheques pré datados, *confirming* e *factoring* entre outros produtos financeiros¹. Por outro lado, os financiamentos de longo prazo referem-se a novos

¹ Ver Alcarva (2011)

projetos e a aquisições de novos ativos que apresentam por norma vidas úteis superiores a um ano (Matias & Esperança, 2009). Neste tipo de maturidade encontram-se produtos como empréstimos de longo prazo bem como o *leasing* financeiro e operacional ou *lease-back*, entre outras.

De acordo com a equação fundamental da contabilidade, o valor do ativo é sempre igual à soma dos passivos com os capitais próprios de uma empresa. A literatura mostra que os indicadores mais utilizados (Tabela 1) para a medição da estrutura de capitais e maturidade da dívida são: a solvabilidade; o endividamento que poderá ser subdividido em endividamento de curto e longo prazo e o rácio de autonomia financeira (Matias & Esperança, 2009).

Tabela 1 – Rácios da Estrutura de Capitais

Rácio	Fórmula
Endividamento	$\frac{\textit{Passivo Total}}{\textit{Ativo Total}}$
Endividamento de Curto Prazo (DCP)	$\frac{\textit{Passivo Corrente}}{\textit{Ativo Total}}$
Endividamento de Longo Prazo (DLP)	$\frac{\textit{Passivo Não Corrente}}{\textit{Ativo Total}}$
Autonomia Financeira	$\frac{\textit{Capital Próprio}}{\textit{Ativo Total}}$

O rácio de endividamento relaciona as dívidas de uma empresa com o valor dos seus ativos. No caso de o rácio ser superior à unidade, estamos perante uma empresa que apresenta capitais próprios negativos. Quanto mais próximo de um, maior é a dependência dos credores, o que aumenta consideravelmente o custo financeiro bem como o risco de falência. O rácio de endividamento pode ser desagregado em endividamento de curto e de longo prazo o que permite verificar o peso que cada componente apresenta junto dos ativos da empresa. Assim, quanto maior for o valor do rácio de curto prazo maior é o risco da empresa (Matias & Esperança, 2009). Neste estudo optou-se pela seleção do endividamento de curto prazo e de longo prazo para estudar a maturidade da dívida dado que abarcam a totalidade da dívida de uma empresa, tal como Hall, Hutchinson, & Michaelas (2000). Vários estudos escolhem a o endividamento total (por exemplo, Vieira & Novo, 2010 e Proença, 2012;), no entanto,

esta variável apresenta-se como uma agregação das anteriores o que não permite evidenciar que fatores estão na base da tomada de decisão de curto e longo prazo.

Por fim, a autonomia financeira expressa a proporção dos ativos que são financiados pelos capitais próprios. Se o rácio se situar perto da unidade as empresas apresentam uma dependência reduzida dos credores o que reduz consequentemente os custos financeiros.

2.2. Caracterização das PME

A Comissão Europeia concebeu em 2003 a definição de PME baseando-se em três critérios: número de colaboradores, volume de negócios e valor dos ativos. Para uma empresa ser incluída em determinada categoria terá de verificar dois dos três critérios (Comissão Europeia, 2003, 2007, 2009). Os critérios e os limites fixados são os constantes da Tabela 2.

Tabela 2 – Enquadramento das PME

Categoria das Empresas	Nº. Trabalhadores	Total do Volume de Negócios		Total de Balanço
Micro	<10	≤ 2 Milhões de euros	ou	≤ 2 Milhões de euros
Pequena	<50	≤ 10 Milhões de euros	ou	≤ 10 Milhões de euros
Média	<250	≤ 50 Milhões de euros	ou	≤ 43 Milhões de euros

Fonte: Comissão Europeia (2003).

O relatório anual das PME (Muller, Gagliardi, Caliandro, Bohn, & Klitou, 2014) relata a importância das PME na Europa nos anos de 2013/2014. Esta importância é por demais evidente, as PME representam cerca de 99,8% do número total de empresa do sector não financeiro existente na União Europeia (UE) composta por 28 países, sendo que 90% destas empresas são micro empresas, ou seja, empregam menos de dez trabalhadores. As PME apresentam-se como a espinha dorsal da Europa, empregando dois em cada três trabalhadores (66,8%) e produzindo cerca de 57,9% do valor acrescentado bruto (Muller et al., 2014). Apesar de serem pouco relevantes individualmente, em termos socioeconómicos, quando observadas na sua globalidade são mais importantes do que as empresas cotadas que apresentam títulos admitidos à cotação (Pontes & Laureano, 2012).

As PME caracterizam-se pela inexistência de separação entre o proprietário e o gestor da empresa, ou seja, na grande maioria das empresas o proprietário é também o gestor. O ambiente em que operam é altamente volátil, o número de clientes e fornecedores é diminuto o que faz com não consigam impor os seus preços (Pontes & Laureano, 2012). Os proprietários integram nas empresas gastos pessoais de forma a amenizar os impostos. Apresentam menor descentralização do poder, maior concentração da propriedade, menor transparência na prestação de informação financeira, reduzidas economias de escala e a imagem da empresa está fortemente ligada à imagem do proprietário (Ang, 1991; Marques & Morgado, 2004).

A responsabilidade limitada é ineficaz porque os proprietários não estão protegidos contra as perdas pessoais resultantes do insucesso das empresas, na medida em que os proprietários dão como garantia os seus bens pessoais para a obtenção de financiamentos (Ang, 1991). No entanto, as PME caracterizam-se por serem mais avessas ao risco (Holmes & Kent, 1991) na medida em que na maioria dos casos a composição da equipa de gestão é imperfeita, ou seja, a mesma carece de profundidade e versatilidade (Ang, 1991), sendo que como na maioria das empresas o proprietário é também o responsável por todas as áreas e o risco não é diversificado por vários investidores (Pontes & Laureano, 2012; Proença, 2012).

2.3. Teorias da estrutura de capitais

Neste subcapítulo pretende-se explicar as principais teorias relacionadas com a estrutura de capitais. Em suma, neste ponto é feita uma breve resenha sobre a abordagem tradicional, a irrelevância da estrutura de capitais, a Teoria do Trade-off, a Teoria da Agência e do *Pecking Order* identificando, assim, o *background* que estará na base da elaboração da investigação empírica.

2.3.1. Abordagem Tradicional

A abordagem tradicional surge com o artigo seminal de Durand (1952) e centra-se no pressuposto que existe uma estrutura ótima de capitais que consegue maximizar o valor das empresas, sendo esta estrutura variável de empresa para empresa.

Em suma, Durand (1952) refere que existe uma estrutura de capitais ótima que permite a maximização do valor da empresa. Esta teoria assenta no facto de o custo do capital próprio ser superior ao do capital alheio pelo que as empresas devem utilizar capital

alheio até ao ponto em que o custo médio ponderado de capital atinja o seu valor mais elevado, e ao atingir tal ponto, as empresas atingem a sua estrutura ótima (Couto & Ferreira, 2009).

2.3.2. Irrelevância da Estrutura de Capitais

A irrelevância nas decisões de financiamento teve a sua génese em 1958 com a publicação do artigo *The Cost of Capital Corporate Finance and Theory of Investment* por parte de Modigliani e Miller. Foi a partir deste momento que se começou a estudar verdadeiramente a importância da estrutura de capitais e, mais concretamente, da maturidade da dívida nas empresas. O impacto desta teoria na literatura financeira foi tão importante que é comparada ao impacto que a Teoria Keynesiana teve na Macroeconomia (Gama, 2000; Weston, 1989).

Modigliani & Miller (1958) desenvolveram uma teoria constituída por três proposições. Esta teoria apresenta um pensamento diferente daquele que até então era divulgado pela abordagem tradicional, ou seja, dentro de determinados pressupostos a estrutura de capitais torna-se irrelevante para o valor das empresas. Apesar da robustez que apresenta na explicação do equilíbrio dos mercados financeiros, peca pela inflexibilidade dos pressupostos assumidos, nomeadamente: a existência de mercados perfeitos e eficientes (princípio da atonicidade); a inexistência de impostos, custos de agência, transação, arbitragem e falência (*financial distress*); os investidores quer sejam pessoas coletivas ou singulares podem investir às mesmas taxas de juro, taxa esta sem risco, o que faz com que todas as entidades apresentem o mesmo risco; a perpetuidade dos fluxos de caixa e o livre acesso à informação.

Da conjugação destes pressupostos, os autores desenvolveram três proposições. A primeira proposição baseia-se no facto de que o valor de mercado de uma empresa é independente da forma como a mesma se financia, assim a forma de como a entidade escolhe a maturidade para a dívida torna-se irrelevante para o valor da empresa (Laureano, 2009). A segunda proposição menciona que o aumento do nível de endividamento de uma empresa aumenta naturalmente a rendibilidade exigida pelos proprietários, ou seja, o custo do capital dos capitais próprios (Mota et al., 2012). Por fim, a terceira proposição baseia-se no facto de que é a política de distribuição de dividendos e não a de financiamento que altera o valor de uma empresa (Mota et al.,

2012). A rentabilidade que um acionista espera receber é igual ou superior à taxa de capitalização que é aplicada pelo mercado às entidades não alavancadas (Dias, 2012).

A teoria cita que à medida que a alavancagem aumenta o custo do capital próprio também aumenta (Mota et al., 2012). É com base nesta premissa que é possível afirmar que a estrutura de capital se mantém inalterada mesmo havendo variações no capital próprio e no capital alheio.

Em 1963, Modigliani & Miller (1963) publicaram um novo trabalho onde fazem referência à importância que os impostos sobre os lucros apresentam na teoria financeira. Esta nova perspectiva aponta para que a estrutura de capital influencie o valor de mercado das entidades, mais concretamente pelo benefício fiscal associado à dedutibilidade dos encargos financeiros. Assim, como em teoria, o capital alheio é “mais barato” do que o capital próprio, o valor ótimo de uma empresa ocorre quando a mesma é financiada exclusivamente por capitais alheios (Mota et al., 2012). Esta teoria apresenta-se como irrealista para que possa ter aderência em termos exequíveis, devido à diversidade dos pressupostos necessários à sua validação (Drobetz & Fix, 2005; Myers, 2001).

Após os estudos de Modigliani e Miller surgiram as seguintes teorias: Teoria do Trade-off, Teoria da Agência e Teoria do *Pecking Order*.

2.3.3. Teoria do Trade-off

A teoria de *Trade-Off* teve a sua origem na década de 70 do século XX, através de autores como Kraus & Litzenberger (1973); Kim (1974) e Scott (1977). Esta teoria revela que as empresas conseguem maximizar o seu valor, recorrendo a financiamento por dívida até ao momento em que o benefício marginal propiciado pela obtenção da poupança fiscal resultante da utilização de dívida igualar o custo marginal de *financial distress* bem como de agência (Kraus & Litzenberger, 1973; Myers, 2001).

Num ambiente onde os juros da dívida são aceites como custo fiscal, o aumento da dívida na estrutura de capitais permite às empresas aumentarem os cash-flows depois de efeitos fiscais para os investidores (Brealey, Myers, & Allen, 2007; Damodaran, 2002). Por um lado, a economia fiscal apresenta-se como favorável para as empresas, por outro lado, o aumento do peso do capital alheio aumenta a probabilidade de as empresas entrarem em situações de pressão financeira, o que no limite pode levar à sua falência

(Myers, 2001). Warner (1977) e Brealey et al. (2007) estudaram os custos relacionados com o aumento da alavancagem e concluíram que estes custos aumentam à medida que o peso da dívida aumenta o que leva ao decréscimo do valor da empresa.

Meggison (1997) revela que as empresas quando entram em falência suportam gastos específicos que as empresas em situação estável não suportam. Kim (1974) e Damodaran (2002) diferenciam estes gastos em diretos e indiretos. Os diretos dizem respeito aos que resultam diretamente da pressão/falência estando relacionados com o pagamento de indivíduos especializados neste tipo de situações. Os indiretos são aqueles intangíveis, ou seja, aqueles que estão relacionados com a perda da capacidade competitiva e que ocorrem quando as empresas estão debilitadas (perda de fornecedores, clientes, colaboradores, apoio das instituições financeiras e demais *stakeholders*).

Em síntese, esta teoria refere que, num mundo onde os mercados de capitais são perfeitos e onde existem custos de falência e impostos, existe para cada entidade uma estrutura ótima de capital próprio e dívida que permite a maximização de valor. Esta teoria revela ainda que a estrutura ótima não poderá ser composta na sua totalidade por dívida devido à existência dos custos de falência (Kim, 1974). A estrutura ótima só é atingida através de uma relação dinâmica entre os benefícios e custos associados à utilização de dívida (Myers, 1984). Os aumentos de capital próprio afastam as empresas do seu valor ótimo pelo que deve ser encarado como uma informação negativa (Proença, 2012; Rasiah & Kim, 2011). A teoria de M&M refere que as empresas encontram-se isoladas do mundo, ao contrário da teoria do *trade-off* que se baseia no facto do mercado de capitais se encontrar em equilíbrio (Kim, 1974).

2.3.4. Teoria da Agência

A Teoria da Agência surgiu com o trabalho de Jensen & Meckling (1976). A problemática da agência pode ser definida como o contrato que se estabelece entre principal e o agente com o intuito de este desenvolver um serviço por sua conta. Este contrato envolve a cedência/delegação de autoridade no agente (Jensen & Meckling, 1976). A teoria da agência refere que gestores e proprietários devem escolher a estrutura de capitais que permita a redução dos custos dos conflitos entre as partes (proprietários, gestores e credores) (Proença, 2012).

As entidades que apresentam elevado volume de *free cash flow* podem deparar-se com conflitos de agência entre os proprietários e os gestores. Os problemas resultam do facto de os proprietários pretenderem ver remunerado o seu capital através da distribuição de dividendos. Por outro lado, os gestores tendem a investir em ativos com maior grau de risco do que os inicialmente acordados e que apresentam um retorno inferior ao expectável pelos proprietários (Drobtetz & Fix, 2005; Laureano, 2009). A dívida permite a redução dos custos de agência ao mitigar os recursos que se encontram disponíveis (Harris & Raviv, 1990; Jensen, 1986; Stulz, 1990).

Por outro lado, se as empresas apresentarem um volume de dívida elevado, podem estar sujeitas aos custos de agência entre os proprietários e os credores. Assim, as empresas devem ser prudentes acerca do volume de dívida na medida em que estes custos tendem a aumentar com o aumento da dívida (Proença, 2012; Rasiah & Kim, 2011).

Myers (1977) alega que as empresas devem utilizar mais dívida quando se encontram numa situação de sobreinvestimento ou mais capital próprio quando estão em situações de subinvestimento.

A teoria do *trade-off* e dos custos de agência entre gestores, proprietários e credores apresentam-se como irrelevantes nas PME na medida em que na maioria das empresas o gestor e o proprietário são a mesma pessoa (Poza & Kishida, 2004; Proença, 2012).

2.3.5. Teoria do Pecking Order

A *Pecking Order Theory* mostra como a volatilidade da informação pode influenciar o financiamento das empresas (Myers, 2001). Myers & Majluf (1984) ao assumirem a existência de assimetria de informação desenvolveram um modelo analítico que demonstra as razões das empresas optarem pelo autofinanciamento em primeiro lugar, e na ausência deste pelo endividamento em vez de recorrerem à emissão de capital próprio.

Fama & French (2002), entre outros autores, apoiam esta teoria afirmando que a alavancagem apresenta uma correlação inversa com a rentabilidade o que verifica a hierarquização das fontes de financiamento. Perante a assimetria de informação, Myers & Majluf (1984) referem que as empresas tendem a financiar-se internamente. No caso de as ações se encontrarem sobreavaliadas os gestores tendem a emitir capital próprio o que poderá significar que o preço das ações vai descer (Myers, 2001).

Ao contrário da teoria preconizada por Modigliani & Miller (1958) esta teoria não vislumbra a existência de uma estrutura de capitais ótima, levando Myers (2001:93) a afirmar o seguinte: «Each firm's debt ratio therefore reflects its cumulative requirement for external financing». Assim, as empresas deverão rever a sua estrutura sempre que se verifiquem desequilíbrios entre as oportunidades de investimento, os *cash flow* gerados e os dividendos (Myers, 2001).

Em síntese, ao contrário da teoria preconizada por Modigliani & Miller (1958) não existe uma estrutura de capitais ótima. Myers (2001:93) revela que «Each firm's debt ratio therefore reflects its cumulative requirement for external financing».

2.4. Fatores determinantes da maturidade da dívida

As teorias anteriormente expostas nasceram das imperfeições existentes nos mercados e com base nas mesmas, vários têm sido os estudos realizados sobre os fatores que influenciam/afetam a estrutura de capitais e mais concretamente a maturidade da dívida das empresas. A literatura aponta para os seguintes fatores: dimensão das empresas; estrutura dos ativos; oportunidades de crescimento; rendibilidade; risco; impostos; alavancagem; *cash-flows* e setor de atividade (p. ex. DeAngelo & Masulis, 1980; Jensen, 1986; Titman & Wessels, 1988; Stulz, 1990; Rajan & Zingales 1995; Guedes & Opler, 1996; Ozkan, 2000, 2002; Myers, 2002; Proença, 2012).

2.4.1. Dimensão das Empresas

A dimensão das empresas e a estrutura dos ativos são dos fatores que mais influenciam a estrutura de capitais (Proença, 2012). Titman & Wessels (1988) evidenciaram empiricamente que a dimensão está positivamente relacionada com o endividamento. Warner (1977) refere que as empresas de maiores dimensões apresentam uma estratégia de negócio diversificada o que lhes permite a estabilidade dos rendimentos reduzindo assim o risco de falência e contribuindo para a liquidação das dívidas. Por outro lado, as PME apresentam uma maior aversão ao risco (Pontes & Laureano, 2012) porque são menos alavancadas preferindo o autofinanciamento, contudo, como foi evidenciado na teoria do *trade-off* a opção por dívida apresenta as suas vantagens (Proença, 2012).

Os estudos existentes que estudam as PME apontam para a existência de uma relação negativa entre a DCP e a dimensão mas positiva com a DLP porque os custos de curto

prazo são geralmente inferiores aos de longo prazo (Esperança et al., 2003; Hall et al., 2000; Michaelas et al., 1999; Proença, 2012). Vieira & Novo (2010), por seu lado, evidenciam a existência de uma relação positiva entre a dimensão e a DLP mas para a DCP não encontraram evidência estatística.

2.4.2. Estrutura dos Ativos

Os ativos de longo prazo de uma empresa podem ser subdivididos em ativos fixos tangíveis e ativos intangíveis. Os primeiros representam os recursos que a empresa detém com carácter de continuidade mas que não tem como destino principal a venda ou transformação no decurso ordinário da atividade, temos como exemplos, os equipamentos, edifícios ou terrenos. Os ativos intangíveis, por sua vez, são ativos que não apresentam substância física, ao contrário dos anteriores. Nesta categoria estão incluídas as marcas, a propriedade intelectual, entre outros.

Scott (1977) e Titman & Wessels (1988) sugerem que as empresas que possuem na sua estrutura ativos tangíveis poderão utilizá-los como garantia para a obtenção de capitais alheios. Assim, as empresas que apresentam mais ativos que possam servir de colateral tem maior facilidade de acesso a financiamentos bancários bem como a outros produtos financeiros diminuindo, desta forma, o risco de incumprimento (Proença, 2012; Vieira & Novo, 2010). Titman & Wessels (1988) e Harris & Raviv (1990) afirmam que as empresas que apresentam mais ativos tangíveis tendem a apresentar um valor mais elevado de liquidação e conseqüentemente maiores níveis de dívida (Proença, 2012).

Morris (1976) afirma que as empresas devem tentar conjugar a maturidade dos seus passivos com a dos seus ativos. A falta de paridade entre a maturidade dos ativos e dos passivos pode ser arriscada para o futuro das empresas na medida em que os *cash-flows* gerados poderão tornar-se insuficientes para o cumprimento do serviço da dívida.

Por um lado, se a maturidade dos ativos for superior à dos passivos as empresas poderão não conseguir obter *cash-flows* para cumprir com as obrigações assumidas. Por outro, se a maturidade dos passivos for superior é também arriscado, visto que é necessário cumprir com as obrigações assumidas mesmo que os ativos não gerem qualquer retorno criando a necessidade de efetuar novos investimentos (Laureano, 2009).

Myers (1977) analisa este fenómeno tendo em conta os problemas de agência. O autor defende que as empresas devem sincronizar o cumprimento da dívida com a redução de

valor dos ativos reduzindo, desta forma, os problemas de agência e do subinvestimento. A conjugação das maturidades permite às entidades dilatar a maturidade da dívida sem aumentar os custos de agência (Laureano, 2009) conseguindo ainda, a redução do risco de liquidez (Morris, 1976; Stohs & Mauer, 1996). Assim, as empresas que apresentem na sua estrutura um peso maior de ativos de médio e longo prazo devem apresentar dívidas com maturidades mais acentuadas.

Os estudos recentes relacionados com o impacto da estrutura de ativos na estrutura de capitais das PME sugerem a existência de uma relação positiva dos ativos tangíveis com a DLP (Hall et al., 2000; Vieira & Novo, 2010). Por outro lado, a relação entre os ativos tangíveis e a DCP não são tão lineares, Hall et al. (2000), Vieira & Novo (2010) e Proença (2012) apontam para a existência de uma relação negativa e Esperança et al. (2003) apontam para o sentido oposto, ou seja, para a existência de uma relação positiva.

2.4.3. Oportunidades de Crescimento

A inclusão de dívida na estrutura de capitais pode ser benéfica para as empresas dada a dedutibilidade dos gastos financeiros. No entanto, o nível ótimo de dívida é influenciado pela existência de oportunidades de crescimento, ou seja, pelas amortizações, depreciações ou provisões (Proença, 2012). Estes elementos caracterizam-se essencialmente pelo facto de as empresas não terem de desembolsar fluxos de caixa para os obter. As amortizações e depreciações referem-se ao desgaste/uso acumulado dos ativos enquanto as provisões referem-se a obrigações legais ou construtivas de montante ou tempestividade incertos.

Ross (1977) refere que a relação expectável é positiva porque quanto maiores forem as oportunidades de crescimento menor são os custos financeiros. Jensen (1986) e Stulz (1990), por seu turno, indicam que a relação positiva deve-se aos problemas de sobreinvestimento, ou seja, o reconhecimento por parte dos credores das oportunidades de crescimento garante às empresas a obtenção de financiamento mais facilmente.

No extremo oposto, Hovakimian et al. (2001) menciona que o crescimento apresenta uma relação negativa com o endividamento porque as empresas preferem financiar o seu crescimento através do autofinanciamento (resultados retidos) em vez do recurso a

dívida, selecionando os problemas de subinvestimento como causa provável (Fama & French, 2002; Proença, 2012; Rajan & Zingales, 1995).

As empresas que apresentam maiores oportunidades de crescimento estão mais sujeitas aos problemas de agência entre os proprietários e os credores apresentando assim uma relação negativa entre o endividamento e o crescimento (Barnea, Haugen, & Senbet, 1980; Myers, 1977). As empresas que apresentam mais oportunidades de crescimento tendem a recorrer menos a dívida do que as empresas que apresentam maiores níveis de ativos tangíveis, esta situação deve-se fundamentalmente ao facto de as oportunidades de crescimento não poderem ser utilizadas como garantia junto dos credores (Myers, 1977, 2002).

Os trabalhos mais recentes apontam para a existência de uma relação positiva entre as oportunidades de crescimento e a maturidade da dívida (Hall et al., 2000; Esperança et al., 2003; Mira & Gracia, 2003 e Proença, 2012).

2.4.4. Rendibilidade

A teoria do *Pecking Order* defende que as empresas apresentam uma hierarquia de fontes de financiamento, em primeiro lugar optam pela utilização dos resultados gerados, caso não sejam suficientes optam pelo recurso a capitais alheios e só depois de esgotadas estas alternativas recorrem à emissão de capitais próprios (Myers, 1984).

A teoria do *Pecking Order* não tem em conta o defendido por Modigliani & Miller (1963), ou seja, dado que as empresas recorrem em primeiro lugar aos resultados gerados não aproveitam as vantagens fiscais decorrentes da dedutibilidade dos gastos financeiros e não enfrentam o risco inerente ao financiamento por dívida (Fama & French, 2002).

Holmes & Kent (1991) referem que nas PME os gestores são na maioria dos casos também os proprietários das empresas pelo que o risco não é diversificado por um vasto conjunto de investidores e os gestores.

Na literatura, a grande maioria dos autores revela a existência de uma relação negativa entre a rendibilidade a maturidade da dívida (Hall et al., 2000; Esperança et al., 2003; Mira & Gracia, 2003 e Proença, 2012).

2.4.5. Risco

O rácio de liquidez permite mensurar a capacidade das empresas em solver as suas dívidas de curto prazo (Proença, 2012). As empresas enfrentam problemas de liquidez quando os *cash-flows* gerados no curto prazo são inferiores ao serviço da dívida (Correia, Brito, & Brandão, 2014).

Diamond (1991) desenvolveu um modelo que relaciona o risco de liquidez e o financiamento de curto prazo. O autor define risco de liquidez como o risco das empresas liquidarem os seus ativos de forma ineficiente dado que não conseguiram o refinanciamento da dívida (Laureano, 2009). Sarkar (1999) defende que o risco de falência e de liquidez poderá influenciar a escolha da maturidade da dívida. Assim, dentro desta visão, o financiamento com maturidades mais longas poderá ser benéfico para as empresas (Laureano, 2009).

As PME caracterizam-se essencialmente por optarem mais por dívida de curto prazo do que de longo prazo (Muller et al., 2014) devido aos problemas de fundo de maneo (Proença, 2012), esta situação leva a que as empresas com baixos rácios de liquidez estejam mais dependentes de dívida de curto prazo. Urbano (2011) concluiu ainda que as entidades que apresentam rácios mais elevados de liquidez preferem financiar-se com maturidades mais longas.

2.4.6. Impostos

Desde o trabalho desenvolvido por Modigliani & Miller (1963) sobre a estrutura de capitais que a influência dos impostos na estrutura de capitais das empresas tem sido alvo de vários estudos, principalmente no que diz respeito a dicotomia entre a dedutibilidade e os benefícios fiscais. O financiamento através do recurso a dívida apresenta o benefício fiscal inerente à dedutibilidade dos juros ao contrário do financiamento através de capitais próprios na medida em que os dividendos não são alvo de dedução. Esta situação permite explicar a preferência por dívida.

Lewis (1990) argumenta que os impostos não apresentam qualquer influência na maturidade da dívida caso sejam definidos em simultâneo os níveis ótimos de alavancagem e de maturidade. O autor refere ainda que os impostos são irrelevantes para a escolha da maturidade da dívida quando os impostos se apresentam como a única imperfeição do mercado.

Os estudos de Ozkan (2000, 2002) e Urbano (2011) apontam para a existência de uma relação positiva mas insignificante em termos estatísticos, concluindo que a taxa de imposto não é um fator determinante na tomada de decisão por parte dos gestores.

2.4.7. Alavancagem

As empresas com níveis mais elevados de alavancagem financiam-se tendencialmente no longo prazo de forma a minimizarem a exposição ao risco de falência (Morris, 1992). Johnson (2003) revela que as empresas que reduzem a maturidade da dívida como forma de mitigação dos problemas de subinvestimento tendem a diminuir a alavancagem de forma a minimizar o risco associado à maturidade de curto prazo (Palma, 2011).

Leland & Toft (1996) desenvolveram um modelo que relaciona o risco de falência e a estrutura ótima de capitais. Os autores concluíram, de acordo com este modelo, que o ponto ótimo da alavancagem depende da maturidade da dívida, sendo maior quando a empresa se financia no longo prazo e menor caso se financie no curto prazo.

Antoniou, Guney, & Paudyal (2006), García-Teruel & Martínez-Solano (2010), Palma (2011) e Stohs & Mauer (1996) apontam para a existência de uma relação positiva entre a alavancagem e a maturidade da dívida nas empresas.

2.4.8. Cash-Flow

As empresas com maior nível de *cash-flow* tendem a utilizar mais dívida como forma de disciplinar os gestores a tomarem as melhores decisões de investimentos em vez de desperdiçarem os *cash-flows* em projetos abaixo do custo de capital (Jensen, 1986; Stulz, 1990).

O estudo das PME espanholas, por parte de Mira & Gracia (2003), e das portuguesas, por parte de Proença (2012), revela a existência de uma relação negativa entre os *cash-flows* e o endividamento, esta situação vem de encontro ao preconizado pela teoria do *Pecking Order*, ou seja, as empresas preferem em primeiro lugar o autofinanciamento e só depois de esgotado é que avançam para a emissão de dívida. Poza & Kishida (2004) referem que dada a insignificância dos problemas de agência nas PME é expectável a existência de uma relação negativa.

2.4.9. Setor de Atividade

Harris & Raviv (1991) mencionam que empresas que operam no mesmo setor de atividade tendem a apresentar características comuns no que diz respeito à estrutura de capitais e mantem os níveis de endividamento ao longo do tempo. Esta situação não se verifica para empresas que operam em diferentes setores de atividade (Proença, 2012).

Lev (1969), Harris & Raviv (1991) e Hall et al. (2000) referem mesmo que o setor de atividade é um dos principais fatores que afeta a estrutura de capitais de uma empresa. Ang (1991) confirma o exposto anteriormente e revela ainda que as PME tendem a utilizar a média do setor para definirem a sua estrutura de capitais na medida em que os gerentes das PME, como já referido anteriormente, não possuem os conhecimentos básicos e necessários de gestão financeira (Pontes & Laureano, 2012; Proença, 2012).

Myers (1984) por seu lado, contradiz totalmente o preconizado anteriormente, referindo que cada empresa apresenta o seu próprio rácio de endividamento, como as PME tendem a operar em nichos de mercado, a média do setor em que atuam tende a ser insignificante. Em suma, o autor afirma que o setor de atividade em que uma empresa opera não tem qualquer influência na escolha da estrutura de capitais (Proença, 2012).

Estudos recentes, realizados por Hall et al. (2000) e Esperança et al. (2003) realçam a importância que o setor de atividade apresenta na estrutura de capitais das PME. Os autores evidenciam estatisticamente a importância que o setor de atividade tem sobre a estrutura de capitais, apontando ainda para a variação dos rácios de dívida de setor para setor (Proença, 2012).

2.5. Estudos Relevantes

O estudo da maturidade da dívida nas PME é relativamente recente dado que o acesso a informação destas empresas é relativamente escasso quando comparado com as grandes entidades cotadas. Na análise das PME de um único país encontramos os trabalhos de Marques & Morgado (2004) nomeadamente no estudo das PME industriais; Vieira & Novo (2010); Proença (2012) e Costa et al. (2014) para as empresas portuguesas. As PME espanholas foram estudadas por López-Gracia & Mestre-Barberá (2011) e Méndez (2013). Michaelas et al. (1999) e Hall et al. (2000) estudaram as PME do Reino Unido. Bhaird & Lucey (2010), por outro lado, optaram pelas PME irlandesas, enquanto as

PME gregas foram analisadas por Daskalakis & Thanou (2010). Hernández-Cánovas & Koëter-Kant (2005) por fim, analisaram 3.366 PME de 19 países europeus.

Os resultados obtidos nos estudos anteriores apontam para as seguintes conclusões: a estrutura dos ativos apresenta uma relação negativa com a DLP e positiva com a DCP (Hall et al., 2000; Proença, 2012; Vieira & Novo, 2010). A dimensão apresenta resultados positivos como a DLP e negativos com a DCP (Esperança et al., 2003; Hall et al., 2000; Michaelas et al., 1999; Proença, 2012; Vieira & Novo, 2010). A rentabilidade, por sua vez, apresenta uma relação negativa com a DCP e DLP (Esperança et al., 2003; Hall et al., 2000; Michaelas et al., 1999; Proença, 2012; Vieira & Novo, 2010).

As oportunidades de crescimento exibem relações positivas com a DCP e DLP, no entanto, os resultados não são estatisticamente significativos (Esperança et al., 2003; Hall et al., 2000; Marques & Morgado, 2004; Michaelas et al., 1999; Proença, 2012; Vieira & Novo, 2010). A liquidação ostenta uma relação positiva com a DLP e negativa com a DCP (Proença, 2012; Urbano, 2011). Marques & Morgado (2004) apontam, ao contrário dos anteriores, para a existência de uma relação não monótona tal como preconizado por Diamond (1991, 1993). O *cash-flow*, por se turno, revela uma relação negativa com a DCP e DLP (Mira & Gracia, 2003; Proença, 2012). Por fim, Costa et al. (2014) encontraram uma relação positiva estatisticamente significativa entre os impostos e a DCP e DLP.

Na Tabela 3 são evidenciadas as relações encontradas nos principais estudos empíricos.

Tabela 3 – Principais estudos empíricos e as relações encontradas

Variável	Evidência Empírica	Relação Esperada
Estrutura dos Ativos	Marques & Morgado (2004) Hall et al. (2000); Proença (2012) e Vieira & Novo (2010)	Positiva (Endividamento Total) Positiva (DLP) e Negativa (DCP)
Dimensão da Empresa	Esperança et al. (2003); Hall et al. (2000); Michaelas et al. (1999); Proença (2012) e Vieira & Novo (2010)	Positiva (DLP) e Negativa (DCP)
Oportunidades de Crescimento	Esperança et al. (2003); Hall et al. (2000); Marques & Morgado (2004); Michaelas et al. (1999); Proença (2012) e Vieira & Novo (2010)	Insignificante Estatisticamente mas apresenta relações positivas com a DCP e DLP

Variável	Evidência Empírica	Relação Esperada
Rendibilidade	Hall et al. (2000); Esperança et al. (2003); Mira & Gracia (2003) e Proença (2012) Antoniou et al. (2006)	Negativa (DLP e DCP) Insignificante Estatisticamente (Endividamento Total)
Liquidez	García-Teruel & Martínez-Solano (2010) Proença (2012) e Urbano (2011) Marques & Morgado (2004)	Convexa (endividamento total) Positiva (DLP) e Negativa (DCP) Não monótona
Cash-flow	Mira & Gracia (2003) e Proença (2012)	Negativa (DCP e DLP)
Taxa Efetiva de Imposto	García-Teruel & Martínez-Solano (2010) e Costa et al. (2014)	Positiva (Endividamento Total)

2.6. Hipóteses de Investigação

Neste estudo pretende-se caracterizar a importância da maturidade da dívida nas PME de Portugal, Espanha, Itália, Grécia, França e Alemanha e conhecer quais os fatores que assumem especial destaque na tomada de decisão por parte dos gestores e se estes fatores são idênticos em todos os países.

Poza & Kishida (2004) referem que os problemas de agência nas PME são praticamente irrelevantes dado que os gestores e os proprietários são na maioria das empresas as mesmas pessoas. Desta forma pretende-se comprovar, não só, se a teoria da agência se aplica nas PME alvo de estudo bem como verificar a importância das teorias do *Pecking Order* e do *Trade-off*. A abordagem tradicional e a irrelevância da estrutura de capitais a inexistência das mesmas não são alvo de estudo.

Titman & Wessels (1988) concluíram que a dimensão das empresas está positivamente relacionada a maturidade da dívida. Desde esta altura, vários têm sido os autores que têm estudado a relação da dimensão com a maturidade da dívida. Os resultados apontam para a existência de uma relação positiva entre a dimensão e a DLP e negativa com a DCP (Michaelas et al., 1999; Hall et al., 2000; Esperança et al., 2003; Vieira & Novo, 2010; Proença, 2012). Desta forma, formulam-se as hipóteses:

H1.a: *A dimensão apresenta uma relação positiva com a DLP*

H1.b: *A dimensão apresenta uma relação negativa com a DCP*

As empresas que apresentam um rácio elevado de ativos que possam servir como garantia na obtenção de financiamentos tendem a optar pela dívida de longo prazo (Myers, 1977; Titman & Wessels, 1988). A literatura aponta para a existência de uma relação positiva entre a DLP e ativos de longo prazo (Michaelas et al., 1999; Hall et al., 2000; Vieira & Novo, 2010 e Proença, 2012). Por outro lado, os resultados obtidos com o estudo da DCP são contraditórios. Contudo, estudos recentes de Vieira & Novo (2010) e Proença (2012) apontam no sentido de que as empresas que apresentam uma maior estrutura ativos tangíveis recorram menos a dívida de curto prazo. Assim, são definidas as hipóteses:

H2.a: *A estrutura dos ativos apresenta uma relação positiva com a DLP*

H2.b: *A estrutura dos ativos apresenta uma relação negativa com a DCP*

As oportunidades de crescimento e a sua relação com o endividamento não apresenta consenso na literatura. Recentemente, alguns estudos apontam no caminho da existência de uma relação positiva com o endividamento (Michaelas et al., 1999; Hall et al., 2000; Esperança et al., 2003; e Proença, 2012). Assim:

H3.a: *As oportunidades de crescimento apresentam uma relação positiva com a DLP*

H3.b: *As oportunidades de crescimento apresenta uma relação positiva com a DCP*

As empresas que apresentam maiores níveis de rendibilidade recorrem menos a financiamentos de longo prazo (Myers & Majluf, 1984). A rendibilidade está intimamente ligada à teoria do *pecking order* que defende a existência de uma relação hierárquica nas fontes de financiamento. Os resultados obtidos para as PME evidenciam a existência de uma relação negativa entre o endividamento e a rendibilidade (Michaelas et al., 1999; Hall et al., 2000; Esperança et al., 2003; Mira & Gracia, 2003; e Proença, 2012). Assim, formulam-se as hipóteses:

H4.a: *A rendibilidade apresenta uma relação negativa com a DLP*

H4.b: *A rendibilidade apresenta uma relação negativa com a DCP*

A literatura evidencia resultados contraditórios no que diz respeito ao impacto do risco de liquidez no endividamento das empresas. Existem estudos que apontam para uma relação negativa com a DCP (Proença, 2012 e Muller et al., 2014) que contrastam com

os que apontam para uma relação positiva com a DLP (Urbano, 2011 e Proença, 2012). Desta forma, formulam-se as seguintes hipóteses:

H5.a: *O risco de liquidez apresenta uma relação positiva com a DLP*

H5.b: *O risco de liquidez apresenta uma relação negativa com a DCP*

Jensen (1986) e Stulz (1990) expõem que as empresas que apresentam maiores níveis de *cash-flow* tendem a utilizar mais dívida. Contudo, Poza & Kishida (2004) alertam para o facto de que nas PME os gestores e proprietários são na maioria das empresas a mesma pessoa o que faz com que os problemas de agência sejam praticamente irrelevantes. Tal como evidenciado por Mira & Gracia (2003) e Proença (2012) é expectável a existência de uma relação negativa entre o endividamento e o *cash-flow*.

H6.a: *O cash-flow apresenta uma relação negativa com a DLP*

H6.b: *O cash-flow apresenta uma relação negativa com a DCP*

O financiamento através do recurso a dívida apresenta o benefício fiscal inerente à dedutibilidade dos juros ao contrário do financiamento através de capitais próprios na medida em que os dividendos não são alvo de dedução. Desta forma, pretende-se verificar se a taxa de imposto apresenta uma relação positiva com a maturidade da dívida.

H7: *A taxa efetiva de imposto relaciona-se positivamente com a maturidade da dívida.*

As PME caracterizam-se por apresentarem empresários que não dispõem de conhecimento e capacidades de gestão financeira que lhes permita seleccionar a melhor estrutura de capitais (Ang, 1991; Pontes & Laureano, 2012). Assim, é importante estudar até que ponto existem diferenças nas tomadas de decisão entre setores de atividade de forma a disponibilizar conhecimento que ajude os empresários a tomarem as melhores decisões. Na literatura encontram-se estudos que apontam para a existência de diferenças entre os setores de atividade (Hall et al., 2000; Esperança et al., 2003 e Proença, 2012).

H8: *O setor de atividade afeta a maturidade da dívida nas PME.*

Capítulo 3 – Descoberta de Conhecimento em Bases de Dados

Neste capítulo realiza-se um breve enquadramento teórico sobre a temática do *Business Intelligence* (BI) e da descoberta de conhecimento em bases de dados (DCBD) e mais detalhadamente de um dos processos da mesma: o DM. O capítulo inicia-se com a definição do conceito de BI seguindo-se a contextualização da DCBD. Após a definição dos conceitos anteriores são expostas as principais aplicações de DM e apresentam-se os principais conceitos, métodos, metodologias e técnicas de DM que podem ser utilizadas no estudo da maturidade da dívida.

3.1. *Business Intelligence*

As novas tecnologias revolucionaram o tratamento da informação, quer na recolha, quer no seu armazenamento, bem como na seleção e consequente distribuição. As empresas devido às condicionantes existentes no mercado em termos de oferta e procura reconheceram a importância do conhecimento organizacional implementando estratégias ao nível da melhoria de processos e produtos bem como ao nível da satisfação dos clientes, integrando as tecnologias da informação (TI) no seio dos sistemas organizacionais (Santos & Ramos, 2009).

A necessidade de desenvolver processos para a extração de conhecimento das bases de dados surge como resposta aos seguintes desafios (Tan, Steinbach, & Kumar, 2006):

Distribuição da informação – é fundamental a redução da quantidade de informação, consolidação dos resultados provenientes de diversas fontes bem como garantir a segurança da mesma quer em termos de privacidade quer em termos de qualidade;

Complexidade e Heterogeneidade – as novas bases de dados agregam não só dados categóricos e numéricos como também fontes complexas de dados com origem em páginas *web* ou em redes sociais, entre outras fontes;

Escala – os algoritmos de análise devem ser preparados para suportar grandes volumes de dados ao contrário das técnicas tradicionais;

Dimensão – o exponencial crescimento da capacidade de processamento e o consequente decréscimo dos custos de tratamento de dados permite desenvolver novas estruturas com um vasto conjunto de informação associada.

O grande benefício dos sistemas de BI para as organizações é a sua enorme capacidade de transformar informação em conhecimento. Esta situação permite às organizações reduzirem custos, aceder a nova informação em tempo útil, facilidade no tratamento dos dados, melhor eficiência na gestão dos recursos disponíveis bem como a otimização dos investimentos em sistemas de informação (Caetano, 2013).

3.2. Processo de Descoberta de Conhecimento em Bases de Dados

As entidades necessitam de tomar decisões de forma rápida, estratégica e segura que lhes permita serem competitivas nos mercados onde desenvolvem as suas atividades. O custo de armazenamento da informação tem vindo a decrescer consideravelmente o que leva as organizações a investir em novas tecnologias. Devido à abundância e a facilidade de acesso aos dados, a área da DCBD é de extrema importância para as organizações (Maimon & Rokach, 2010).

A DCBD tem como finalidade criar/desenvolver técnicas que a partir das grandes bases de dados consigam identificar padrões e tendências com o intuito de produzir novo conhecimento tendo como critérios de validação (terão de ser aplicáveis a novos dados) o interesse e a inovação (Fayyad, G Piatetsky-Shapiro, et al., 1996).

O processo DCBD apresenta-se como um «processo não trivial de identificação de padrões válidos e potencialmente úteis, perceptíveis a partir dos dados» (Cios, Pedrycz, Swiniarski, & Kurgan, 2007: 10; Fayyad et al., 1996). De acordo com Santos & Ramos (2009) esta definição pode ser decomposta em vários pontos, ou seja, **padrão** caracteriza-se pelos modelos ou relações nos dados que devem ser **perceptíveis**, caso não sejam imediatamente visíveis devem-no ser após determinado período. Os **dados** consubstanciam-se na informação armazenada nas *Data Warehouses*. O **processo** está dependente da realização dos diferentes passos iterativos, que vão desde a seleção dos dados até à interpretação dos resultados obtidos. O **não trivial** advém do facto de poder compreender a pesquisa de estruturas, modelos ou padrões. O processo além de iterativo é também **interativo** porque requer a participação do utilizador sempre que seja necessário tomar decisões.

O processo de DCBD inclui diversas fases, sendo a fase do DM a mais importante, sendo nesta fase que se procura encontrar relações para os dados. Os termos DM e DCBD são frequentemente utilizados com o mesmo significado, no entanto, o processo

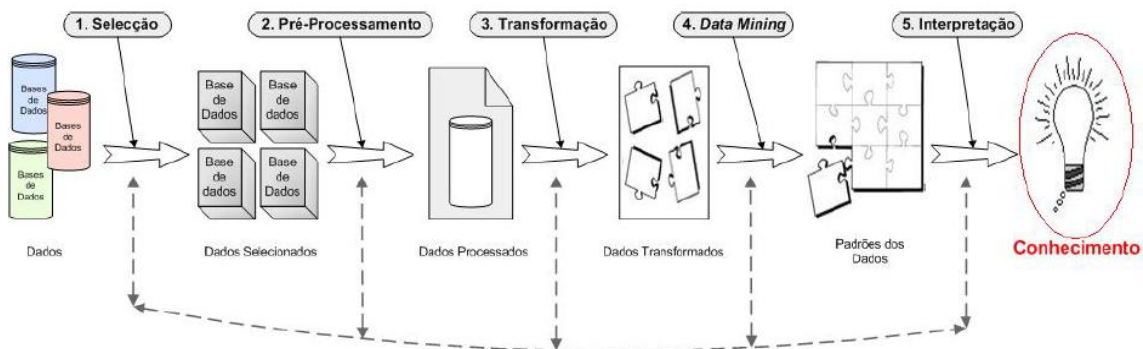
de DM é apenas um dos passos da DCBD (Han & Kamber, 2001). Assim, e ao longo do próximo capítulo, é utilizada uma noção em sentido lato.

3.2.1. Fases do Processo de DCBD

As diversas fases do processo de DCBD podem ser visualizadas na

Figura 1, onde se incluem a seleção, pré-processamento, tratamento e transformação dos dados, DM e a interpretação dos resultados.

Figura 1 – Processo de DCBD



Fonte: Portelinha (2005: 12)

Cada uma das fases pode ser resumida (Santos & Ramos, 2009):

Seleção dos Dados – Após definir-se o âmbito e os objetivos do processo, procede-se à recolha dos dados que apresentem características úteis;

Tratamento e Pré-processamento dos Dados – tratamento das omissões e incorreções nos dados, problemas de distribuição dos dados, entre outros. Nesta fase procede-se à limpeza e verificação dos dados selecionados anteriormente;

Transformação – procura-se identificar as características mais importantes nos dados de forma a transformar, reduzir ou modificar as variáveis;

Data Mining – fase onde se selecionam os métodos e técnicas para a extração de padrões e tendências nos dados;

Interpretação e Avaliação – visualização do conhecimento extraído da fase anterior no sentido de proceder à sua avaliação e interpretação, sendo ainda possível retornar às fases anteriores para uma nova iteração.

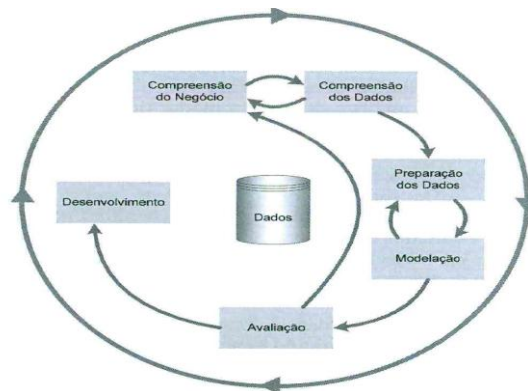
3.3. Metodologias do *Data Mining*

As principais metodologias de DM são a CRISP – DM (*Cross Industry Standard Process for Data Mining*) e SEMMA (*Sample, Explore, Modify, Model, Assessment*). O CRISP – DM foi desenvolvido em 1996 por três analistas que representavam a NCR Systems Engineering Copenhagen, a DaimlerChrysler AG e a SPSS Inc. Por outro lado, a SEMMA foi construída e desenvolvida pela empresa SAS. No entanto, a CRISP – DM é a metodologia com maior aceitação em termos mundiais pelo que é abordada detalhadamente no próximo ponto.

3.3.1. Metodologia CRISP-DM

O CRISP-DM é descrito por Chapman et al. (2000) como um modelo que retrata o ciclo de vida dos projetos de DM. A Figura 2 ilustra as seis fases essenciais desta metodologia que, à semelhança da DCBD, permite que se retroceda a etapas anteriores de forma a alterar-se ou incluir-se novos dados sendo posteriormente retomado o processo (Santos & Ramos, 2009). Cada fase integra diversas tarefas que permitem tornar o processo perceptível para os diversos utilizadores (Cios et al., 2007; Larose, 2005).

Figura 2 – Metodologia CRIP – DM



Fonte: Santos & Ramos (2009: 109)

As fases da metodologia CRISP – DM são sumariadas de seguida (Chapman et al., 2000; Santos & Ramos, 2009):

Compreensão do Negócio – Nesta fase procede-se à compreensão dos objetivos do ponto de vista do negócio, definição da problemática de DM e à criação dum plano preliminar.

Compreensão dos dados – Nesta fase executam-se diferentes tarefas de exploração com vista à compreensão dos dados, identificação de problemas ou subconjuntos relevantes, os quais são posteriormente analisados de modo a identificar-se o conhecimento implícito nos mesmos.

Preparação dos dados – Procede-se à obtenção do conjunto de dados que são alvo de análise pelos algoritmos de DM. As tarefas incluídas nesta fase passam pela seleção, limpeza, construção, identificação e formatação dos dados. Estas tarefas podem ser executadas diversas vezes não existindo uma ordem para a execução das mesmas.

Modelação – Fase onde se selecionam os algoritmos de DM que são aplicados aos dados extraídos da fase anterior. Inclui a seleção das técnicas de modelação, definição dos conjuntos de teste e treino bem como a construção e análise dos modelos.

Avaliação – Os modelos apurados, e que apresentam melhores níveis de qualidade (medido pela confiança estatística, entre outros), são interpretados nesta fase. Com a análise dos modelos pretende-se verificar se os mesmos verificam os objetivos delineados na fase da compreensão do negócio. No término da fase procede-se à decisão da utilização ou não dos modelos. Nesta fase avaliam-se os resultados, revê-se o processo e identificam-se os passos seguintes bem como os resultados esperados.

Desenvolvimento – A identificação dos modelos não pressupõe a conclusão do projeto. A informação recolhida terá de ser organizada e apresentada para que seja útil à organização. Nesta fase incluem-se o planeamento do desenvolvimento, monitorização e manutenção, produção do relatório final assim como a revisão do projeto.

3.4. Data Mining

O termo *data mining* em sentido lato pode ser definido como a «extração de conhecimento de grandes quantidades de dados» (Han & Kamber, 2001).

A definição de *DM* varia de autor para autor tendo em conta as várias áreas do conhecimento. De seguida são evidenciadas algumas das principais definições. Fayyad, Piatetsky-Shapiro, & Smyth (1996: 39–43) definem DM como «a aplicação de algoritmos específicos na extração de padrões dos dados. A maioria dos métodos de DM é baseada em técnicas testadas de *machine learning*, reconhecimento de padrões e estatísticas: classificação, *clustering*, regressão entre outras». Xu, Li, & Zhang (2011: 5) referem-se ao DM como a «extração de conhecimento informativo de uma vasta quantidade de dados, que podem estar expressos de diversas formas. Não interessa o

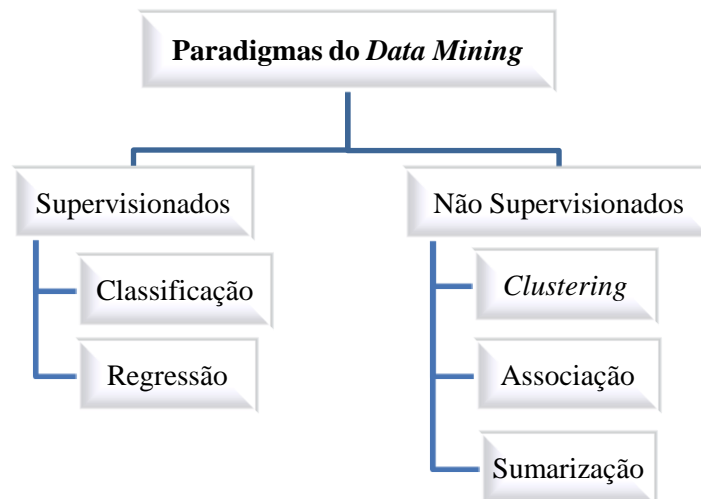
tipo de dados, o principal objetivo do DM é descobrir conhecimento oculto e invisível normalmente na forma de padrões a partir dos repositórios de dados». Por sua vez, Berry & Linoff (2004:7) interpretam como «o processo de exploração e análise de grandes quantidades de dados de forma a encontrar padrões e regras significativas».

Em suma, da análise das definições precedentes, o processo de DM consiste na aplicação de técnicas e métodos a grandes quantidades de dados de forma a descobrir padrões e/ou tendências úteis e relevantes com o intuito de extrair/produzir novo conhecimento.

3.4.1. Paradigmas e tarefas de Data Mining

Os paradigmas são divididos em dois grupos, aprendizagem não supervisionada e aprendizagem supervisionada, conforme visualizado na Figura 3.

Figura 3 – Paradigmas do Data Mining



A. Aprendizagem Não Supervisionada

Neste paradigma não existem variáveis alvo, ou seja, para cada tarefa existem apenas valores de entrada. Os algoritmos de aprendizagem procuram estruturas nos dados que permitam agrupá-los em diferentes classes (Larose, 2005; Maimon & Rokach, 2010; Witten, Frank, & Hall, 2011).

Clustering ou Segmentação

O principal objetivo é agrupar os dados em grupos homogêneos entre si. Devem ser identificados agrupamentos nos dados que apresentem características idênticas. Podem ser estabelecidos, previamente, o número de clusters a formar ou então deixar o

algoritmo definir automaticamente. Os algoritmos de DM mais utilizados são o *K-means* e as redes neuronais do tipo *Kohonen* (Kohonen, 1982), também reconhecidas como mapas auto-organizáveis. Como exemplos desta tarefa temos os estudos de mercado, sendo o principal objetivo a identificação de perfis de clientes homogêneos. Esta separação permite aos utilizadores detetar as diferenças nos dados e identificar novas oportunidades (Santos & Ramos, 2009).

Associação

Esta tarefa é responsável por construir modelos que descrevam dependências significativas entre as variáveis. As regras de associação apresentam grande importância em DM porque identificam atributos direta ou indiretamente relacionados entre si. Trata-se de um objetivo que é muito utilizado na análise dos comportamentos de clientes, com a utilização de regras do tipo: se o cliente compra o produto x também compra o produto y. O algoritmo mais utilizado é o *Apriori* (Agrawal & Srikant, 1994).

Sumarização

Esta tarefa, mais do que tentar prever ou classificar, tem como objetivo explorar os dados, ou seja, pretende descobrir uma explicação para um determinado conjunto de dados, com o intuito de realizar análises exploratórias e descritivas bem como a criação de relatórios. Esta tarefa pode ser utilizada nas empresas, por exemplo, na comparação das vendas entre filiais (Santos & Ramos, 2009).

B. Aprendizagem Supervisionada

A maioria das tarefas de extração de conhecimento de bases de dados apresenta uma aprendizagem supervisionada, isto quer dizer que apresentam uma variável alvo pré-definida. Os métodos supervisionados são aplicados na descoberta da relação entre os dados de entrada (independentes) e os dados de saída (dependentes). A relação descoberta é representada através de um modelo. Comumente, os modelos descrevem e explicam fenómenos que estão ocultos nos dados, podendo ser utilizados na classificação ou previsão dos valores de saída tendo como ponto de partida os valores de entrada (Larose, 2005; Maimon & Rokach, 2010).

Classificação

A tarefa de classificação significa examinar a informação e enquadrá-la em determinado conjunto com características semelhantes. Esta tarefa classifica os dados, construindo o modelo baseando-se num conjunto de treino e nos valores do atributo classificador, para assim determinar a que classes pertencem os novos dados. A classificação é um dos objetivos mais comuns em DM, sendo as árvores de decisão, a indução de regras, as regressões logísticas e as redes neuronais artificiais as técnicas mais utilizadas (Chapman et al., 2000). Em termos de aplicações práticas para a classificação encontram-se a identificação do *rating* das empresas (Bose & Mahapatra, 2001) ou o comportamento dos utilizadores de cartões de crédito (Desai, Crook, & Overstreet, 1996).

Regressão

Pode ser definida como uma tarefa que prevê um valor baseando-se em determinadas variáveis. Esta tarefa tem como principal objetivo encontrar uma função que permita prever uma variável dependente, ou seja, consiste na construção de uma função que represente de forma adequada o comportamento de um conjunto de variáveis independentes. No caso em estudo, utilizam-se algoritmos de predição, ou seja, pretende-se prever quais as características que mais afetam a maturidade da dívida nas empresas. Os algoritmos mais aplicados em problemas de predição são a regressão linear múltipla, as árvores de decisão (por exemplo, com o algoritmo CART), as redes neuronais artificiais e mais recentemente as máquinas de vetores de suporte (Hastie et al., 2009). Na predição encontram-se como exemplos a previsão dos valores de venda de ações na bolsa ou a previsão de preços de imóveis (Silva, 2007).

Na Tabela 4 apresentam-se algumas das técnicas de DM e o fim a que se destinam.

Tabela 4 – Técnicas de DM e o tipo de problemas que solucionam

Técnica de <i>Data Mining</i>	Tipos de Problemas	
	Classificação	Regressão
Árvores de Decisão	X	X
Regressão Linear Múltipla		X
Regressão Logística	X	
Redes Neuronais Artificiais	X	X
Máquina de Vetores de Suporte	X	X
Regras de Classificação	X	

Capítulo 4 – Metodologia

Os vários modelos preditivos que são expostos neste capítulo visam fundamentalmente a extração de padrões e tendências nos dados e a consequente descoberta de novo conhecimento sobre a maturidade da dívida nas empresas. O objetivo é identificar os fatores que mais influenciam a maturidade e se afetam de igual forma os países selecionados.

Devido à flexibilidade e iteração no desenvolvimento das diferentes fases, optou-se pela escolha da metodologia CRISP-DM, na medida em que esta fornece um guião que auxilia no desenvolvimento de projetos de DM. De seguida são evidenciadas as seis etapas do CRISP-DM e as tarefas realizadas em cada uma delas:

4.1. Compreensão do problema

O estudo da maturidade da dívida tem vindo a crescer na literatura. No entanto, não tem sido devidamente abordado, na medida em que grande parte dos estudos privilegia a análise de empresas cotadas, países isolados, ou acontecimentos marcantes. Esta situação fez com que se optasse pelo estudo das PME de seis países pertencentes à UE, mais concretamente Portugal, Espanha, Itália, Grécia, França e Alemanha. Estes países foram selecionados porque pretende-se avaliar se a maturidade da dívida das empresas dos países mais afetados pela crise (Portugal, Espanha, Grécia e Itália) é idêntica à dos países menos afetados (França e Alemanha) e quais os fatores que mais influenciaram a tomada de decisão por parte dos gestores no ano de 2013.

As tomadas de decisões relacionadas com o financiamento são de extrema importância para a sobrevivência das empresas, incidem essencialmente na dicotomia capital próprio ou dívida e qual a maturidade que melhor representa os objetivos das empresas e que permita a maximização do seu valor. A maturidade da dívida teve o seu prólogo na década de cinquenta do século XX com a abordagem tradicional e a irrelevância da estrutura de capitais. A partir destes estudos foram-se diversificando as teorias na literatura, destacando-se as teorias do *Trade-off*, *Agência* e *Pecking Order*.

Em termos de critérios de sucesso a escolha incide sobre o modelo de DM que melhores resultados produzir, ou seja, que seja capaz de prever melhor os valores de maturidade da dívida.

4.2. Compreensão e Preparação dos dados

A amostra foi extraída da Base de Dados da Amadeus que está disponível *online*² e inclui empresas não financeiras de Portugal, Espanha, Grécia, França, Alemanha e Itália selecionadas de acordo com a definição de PME preconizada pela Comissão Europeia. A amostra selecionada inclui 709.752 empresas sendo os dados recolhidos essencialmente de cariz financeiro, provenientes da demonstração da posição financeira e da demonstração de resultados, bem como de rácios financeiros (Anexo 1). Apesar da base de dados Amadeus dispor de informação de muitos países optou-se pela seleção dos anteriores dados, uma vez que o foco principal é analisar se as empresas dos países que foram mais afetadas pela crise das dívidas soberanas apresentam os mesmos índices de maturidade da dívida que as dos países menos afetados.

Na fase da exploração dos dados procedeu-se à análise da informação recolhida de forma a removerem-se as observações que não adicionavam valor bem como das redundantes e duplicadas. Das empresas recolhidas foi de seguida extraída uma amostra aleatória que engloba 62.018 micro empresas (72%), 12.658 pequenas empresas (15%) e 11.482 médias empresas (13%) o que perfaz um total de 86.158 empresas. Em termos de distribuição por país, a amostra é constituída por 26.548 empresas italianas (31%), 11.781 portuguesas (14%), 18.964 espanholas (22%), 15.208 francesas (17%), 7.565 alemãs (9%) e 6.092 gregas (7%).

Comparando a amostra utilizada com as amostras de estudos anteriores (Tabela 5), conclui-se que alguns dos estudos focam-se em empresas cotadas dada a facilidade de acesso à informação bem como em empresas de um determinado país. Assim, a amostra selecionada apresenta a vantagem de incluir dados dos diferentes tipos de PME e que cobrem grande parte dos setores de atividade. Os estudos anteriores apresentam amostras relativamente reduzidas. Desta forma, a presente amostra permite ver até que ponto as diferentes determinantes se aplicam num conjunto mais amplo de observações.

² <https://amadeus.bvdinfo.com> (acessível mediante subscrição)

Tabela 5 – Principais estudos da maturidade da dívida e dimensão da amostra

Autor	Amostra	Período	Fonte	Países
Michaelas et al. (1999)	35000 PME	1986-1995	Lotus One Source Database	Reino Unido
Hernández-cánovas & Koëter-kant (2005)	7669 PME	2000	Entrevistas	19 Países da Europa
Vieira & Novo (2010)	5077 PME	2000-2005	Sistema de Análise de Balanços Ibéricos (SABI)	Portugal
Urbano (2011)	855 Empresas Cotadas	2001-2010	Bloomberg	Grécia, Portugal, Irlanda, Itália e Espanha
Palma (2011)	7724 PME e Cotadas	2002-2009	Amadeus	Portugal e Espanha
Correia (2012)	3524 Empresas Cotadas	2000-2011	Bolsa de Valores	12 Países da UE e Reino Unido
Proença (2012)	12857 PME	2007-2010	Amadeus	Portugal
Costa, Laureano, & Laureano (2014)	1998 PME	2009-2011	Informa D&B	Portugal

Assim, de acordo com as informações obtidas das fases anteriores conclui-se que os recursos existentes são adequados para a execução da investigação empírica.

Na fase da preparação dos dados procedeu-se ainda à identificação de eventuais erros ou de valores omissos. Nos modelos de previsão existem diferentes formas de tratar os valores desconhecidos. Em tarefas de previsão, os algoritmos tendem a não apresentarem resultados fidedignos quando os dados apresentam valores desconhecidos. Caso se procedesse à substituição dos valores poder-se-ia desvirtuar o processo de predição o que não permitiria a extração de conhecimento útil e relevante. Outro dos problemas habitualmente identificados é o que fazer com os valores extremos e com os *outliers*. Em suma, e de forma a normalizar a amostra, procedeu-se à eliminação dos valores omissos, dos valores discrepantes, dos extremos e dos *outliers*³. Esta análise foi efetuada com recurso ao *software* IBM SPSS Modeler sendo complementada com a ferramenta *Excel*.

³ Os *outliers* consistem em observações que são inconsistentes ou que apresentam um grande afastamento das restantes. Os extremos apresentam-se como os pontos máximos e mínimo a partir dos quais se identificam os *outliers* (Pestana & Gageiro, 2005).

A fase da transformação/construção das variáveis só pode ser desenvolvida após a limpeza dos dados. Nesta fase, e de acordo como os elementos recolhidos (Anexo 1) procedeu-se à criação das variáveis (Anexo 2) que integram o presente estudo empírico.

Indicadores da Maturidade da Dívida

Foram selecionadas duas variáveis dependentes para medir maturidade da dívida das empresas, a saber: maturidade de curto prazo (DCP) e maturidade de longo prazo (DLP). Optou-se pela não inclusão do endividamento total como variável porque é uma agregação das duas variáveis selecionadas o que não permitiria verificar os fatores mais importantes na tomada de decisão no curto e longo prazo.

A Tabela 6 mostra as duas variáveis dependentes selecionadas que se baseiam em valores contabilísticos.

Tabela 6 – Indicadores de Maturidade da Dívida

Indicador	Notação	Fórmula
Maturidade de Longo Prazo	DLP	$MLP = \frac{Dívida\ de\ Longo\ Prazo}{Ativo\ Total}$
Maturidade de Curto Prazo	DCP	$MCP = \frac{Dívida\ de\ Curto\ Prazo}{Ativo\ Total}$

De forma a testar as hipóteses de investigação expostas e fundamentadas na revisão da literatura, são utilizadas as seguintes variáveis independentes, classificadas em duas dimensões: características das empresas e variáveis de controlo.

Fatores Determinantes da Maturidade da Dívida

As características das empresas são quantitativas que derivam de elementos provenientes das demonstrações financeiras das empresas. As variáveis selecionadas são utilizadas pelos utilizadores das demonstrações financeiras e oscilam mediante a performance de cada entidade.

Estrutura dos Ativos – As empresas que apresentam um rácio elevado de ativos que possam servir como garantia na obtenção de financiamentos tendem a optar pela dívida de longo prazo (Myers, 1977; Scott, 1977; Titman & Wessels, 1988). A estrutura dos ativos é calculada pela divisão dos ativos fixos tangíveis e intangíveis pelo ativo total.

Oportunidades de Crescimento – Os estudos anteriores não apresentam consenso na relação entre as oportunidades de crescimento e a maturidade da dívida. Michaelas et al. (1999) e Vieira & Novo (2010) utilizaram a taxa de crescimento dos ativos. Hall et al. (2000) utilizou a taxa de crescimento das vendas. Neste estudo optou-se pelo quociente entre as depreciações e amortizações e o ativo total, à semelhança de Correia (2012).

Dimensão – Tal como mencionado anteriormente esta é uma das principais variáveis no que concerne ao estudo da maturidade da dívida. Na literatura existem duas formas de definir a dimensão, pelo logaritmo natural do volume de negócios e mais recentemente, pelo logaritmo natural do ativo (Esperança et al., 2003; Mira & Gracia, 2003 e Proença, 2012). Desta forma utilizar-se-ão as duas abordagens no sentido de se perceber se apresentam o mesmo poder explicativo junto da maturidade da dívida.

Rendibilidade – A literatura aponta para a existência de uma relação negativa entre a rendibilidade e a maturidade da dívida. Dada a variedade de variáveis existentes na literatura optou-se pela inclusão de três variáveis para definir a rendibilidade. Optou-se pela rendibilidade operacional (Myers & Majluf, 1984) e líquida (Hall et al., 2000; Esperança et al., 2003; Mira & Gracia, 2003 e Proença, 2012) do ativo e do capital próprio.

Alavancagem – As empresas que apresentam elevados níveis de endividamento tendem a mitigar a sua exposição através do aumento da maturidade. As empresas que apresentam elevados níveis de dívida de curto prazo como forma de mitigação dos problemas de subinvestimento tendem a reduzir a alavancagem de forma a reduzir a exposição ao risco (Johnson, 2003; Leland & Toft, 1996; Morris, 1992). Palma (2011) aponta para a existência de uma relação positiva entre a alavancagem e a maturidade da dívida.

Risco – Os resultados obtidos em estudos anteriores não são consensuais. As empresas com menor risco financiam-se no curto prazo. No extremo oposto, as empresas que apresentam maior risco financiam-se no longo prazo de forma a evitarem problemas de refinanciamento. Para definir utilizou-se rácio entre o ativo e o passivo corrente (Mira & Gracia, 2003; Urbano, 2011 e Proença, 2012) e o *Z score* (Correia, 2012; Marques & Morgado, 2004; Méndez, 2013).

Taxa de imposto – Os financiamentos permitem a obtenção de vantagens fiscais resultantes da dedutibilidade dos juros (Brick & Ravid, 1985). Assim, pretende-se verificar até que ponto a taxa efetiva de imposto influencia a escolha da maturidade da dívida nas empresas. Esta variável é medida pelo rácio entre o imposto e o resultado antes de imposto (Correia, 2012; Méndez, 2013).

Cash-Flow – As empresas com maior nível de cash-flow utilizam mais dívida de forma a reduzirem os problemas de agência (Jensen, 1986; Stulz, 1990). Poza & Kishida (2004) referem que dada a insignificância dos problemas de agência nas PME é expectável a existência de uma relação negativa. Estas conclusões foram sugeridas também por Mira & Gracia (2003) e Proença (2012).

A Tabela 7 apresenta as diferentes variáveis independentes (características das empresas) e a sua expressão de cálculo.

Tabela 7 – Fatores Determinantes da maturidade da dívida

Indicador	Notação	Fórmula
Estrutura dos Ativos	EST	$EST = \frac{\text{Ativos Fixos} + \text{Intangíveis}}{\text{Ativo Total}}$
Oportunidades de Crescimento	OC	$OC = \frac{\text{Depreciações} + \text{Amortizações}}{\text{Ativo Total}}$
Dimensão da empresa em termos de ativo	DIMA	$DIMA = \text{Logaritmo Natural do Ativo Total}$
Dimensão da empresa em termos de volume de negócios	DIMVN	$DIMVN = \text{Logaritmo Natural do Volume de Negócios}$
Rendibilidade do ativo	ROA	$ROA = \frac{\text{Resultado Líquido}}{\text{Ativo Total}}$
Rendibilidade do capital próprio	ROE	$ROE = \frac{\text{Resultado Líquido}}{\text{Capital Próprio}}$
Alavancagem	ALV	$ALV = \frac{\text{Capital Próprio}}{\text{Passivo Total}}$
Liquidez	LIQ	$LIQ = \frac{\text{Ativo Corrente}}{\text{Passivo Corrente}}$

Indicador	Notação	Fórmula
Z Score - Medida de Avaliação do Risco	Z	$Z = 0,717 * \frac{(Ativo Corrente - Passivo Corrente)}{Ativo Total} + 0,847$ $* \frac{Resultados Transitados}{Ativo Total} + 3,107$ $* \frac{Resultados Antes de Imposto}{Ativo Total} + 0,420$ $* \frac{Capital Próprio}{Passivo} + 0,998 * \frac{Vendas}{Total Ativo}$
Rendibilidade da atividade operacional	RENT	$RENT = \frac{R.A.J.I}{Ativo Total}$
Taxa Efetiva de Imposto	TEI	$TEI = \frac{Imposto}{Resultado antes de imposto}$
Cash-Flow	CFLOW	$CFLOW = \text{Logaritmo Natural do cash Flow}$

Variáveis de Controlo

As variáveis de controlo caracterizam-se por serem tendencialmente qualitativas e imutáveis. Apesar de a literatura não incluir variáveis de controlo optou-se pela sua inclusão dada a significância que as mesmas podem apresentar para o estudo da maturidade da dívida. Assim, seleccionaram-se as seguintes quatro variáveis:

Critério PME (UE): Esta variável assume o valor de micro, pequena e média empresa.

Sector de Atividade: o sector de atividade é decomposto em 16 sectores (Anexo 3).

País: A variável país é composta por: Portugal, Espanha, França, Grécia, Itália e Alemanha.

Caracterização das variáveis em estudo

O conhecimento do comportamento das variáveis em estudo e também da relação existente entre elas ajuda a uma melhor compreensão do problema. Para o efeito são apresentadas as medidas de estatística descritiva mais frequentes e na análise de relações entre duas variáveis recorre-se às medidas de associação/correlação. No caso da relação entre os indicadores de maturidade e o tipo de empresas (micro, pequena e média) e o sector de atividade recorre-se ao teste *one-way* ANOVA (F-test).

A Tabela 8 apresenta as medidas de estatística descritiva das variáveis dependentes da amostra completa. A maturidade de curto prazo apresenta um valor médio de 0,50, ou seja, nas empresas seleccionadas o valor do passivo corrente representa cerca de metade

do valor dos ativos. A mediana, que divide a amostra ao meio, apresenta um valor ligeiramente inferior (0,48). A DCP apresenta um valor mínimo de 0 e um valor máximo de 1,91, ou seja, existem empresas que apresentam capitais próprios negativos, daí o passivo ser superior ao valor do ativo, o que faz com que a variável apresente muita dispersão (DP = 0,26). Por seu lado, a DLP apresenta um valor médio de 0,16. No entanto, metades das observações são iguais ou inferiores a 0,09. Pelo percentil 75 é possível observar que três quartos da amostra apresentam valores iguais ou inferiores a 0,24

Tabela 8 – Estatística Descritiva dos Indicadores da Maturidade

	M⁴	DP	Mín	P 05	P 25	Mediana	P 75	P 95	Máx
DCP	0,50	0,26	0,00	0,10	0,29	0,48	0,69	0,91	1,91
DLP	0,16	0,19	0,00	0,00	0,01	0,09	0,24	0,55	1,92

N = 86.158

A análise da relação entre os dois indicadores evidencia a existência de uma correlação negativa e fraca entre a maturidade de longo e de curto prazo ($r(86.156) = - 0,38, p < 0,05$).

A Tabela 9 apresenta a estatística descritiva dos fatores determinantes da maturidade da dívida da amostra selecionada no ano de 2013. Os ativos que poderão servir de garantia (EST) na obtenção de dívida apresentam um peso médio de 0,20 em relação ao ativo total, sendo que a mediana é de cerca de 0,12. A variável OC que representa o peso das depreciações no ativo sugere que em média as empresas apresentam um valor de 0,04, ou seja, as depreciações apresentam um peso relativamente baixo nas empresas selecionadas. A dimensão medida pelo ativo apresenta um valor médio de aproximadamente 13,46 enquanto se for medida pelo volume de negócios apresenta um valor de 13,59. Estas duas variáveis apresentam uma grande amplitude de variação, ou seja, a primeira varia entre os 7,19 e os 20,90 enquanto a segunda oscila entre o 0 e os 17,73. Esta amplitude faz com que a dispersão dos valores seja também elevada.

Em termos da rentabilidade, a do ativo (ROA) situa-se nos 0,06 com um valor mediano de 0,03, por sua vez, a rentabilidade do capital próprio (ROE) apresenta uma amplitude que vai desde os - 27,87, ou seja, existem empresas que apresentam resultado líquido negativo e os 30,04, situação onde o valor do resultado líquido é cerca de 30 vezes

⁴ M – Média; DP – Desvio Padrão; Mín – Mínimo; P 05 – Percentil 5%; Máx – Máximo

superior ao valor dos capitais próprios. No entanto, se visualizarmos a amostra entre o percentil 05 e o percentil 95 é possível observar que os valores oscilam entre os -0,03 e os 0,73. Ao nível da rentabilidade operacional (RENT) representa em média 0,12 do valor dos ativos das empresas sendo a mediana de 0,10. Em 95% da amostra os valores da rentabilidade operacional situa-se abaixo ou igual a 0,32. Sendo que cerca de 5% apresenta rentabilidades anormais.

A ALV apresenta um valor médio de 1,06 o que significa que em média as empresas apresentam níveis de capitais próprios ligeiramente superiores ao valor dos seus passivos. Contudo, o valor da mediana contradiz esta afirmação na medida em que apresenta um valor de 0,47, ou seja, aproximadamente mais de 45% das observações apresenta um valor de passivo superior ao dos capitais próprios. Na liquidação (LIQ), o ativo corrente é, em termos médios, duas vezes superior ao valor do passivo corrente (M = 2,11). Esta variável apresenta também uma grande dispersão (DP = 2,58) na medida em que oscila entre os 0,0001 e os 75,55. O *Z Score*, que mede o risco das empresas encontra-se na zona cinzenta (M = 2,50). Apenas cerca de 25% amostra se encontra na zona segura (acima dos 2,9). A taxa de imposto (TEI) apresenta um valor médio e mediano relativamente próximos (M = 0,24 contra a Me = 0,23). A amplitude vai desde os - 1,98 e os 1,86 sendo que 90% da amostra está compreendida entre os 0 e os 0,65. Por fim, as empresas apresentam um valor médio do logaritmo do *cash-flow* livre na ordem dos 10,67, variando entre os 1,61 de mínimo e os 17,89 de máximo sendo o desvio padrão de 1,9.

Tabela 9 – Estatística Descritiva dos Fatores Determinantes da Maturidade da Dívida

Variável	M	DP	Mín	P 05	P 25	Mediana	P 75	P 95	Máx
EST	0,20	0,22	0,00	0,00	0,04	0,12	0,31	0,68	1,00
OC	0,04	0,04	0,00	0,00	0,01	0,03	0,05	0,11	0,94
DIMA	13,46	1,83	7,19	10,77	12,08	13,25	14,68	16,79	20,90
DIMVN	13,59	1,84	0,00	10,97	12,22	13,45	14,77	16,99	17,73
ROA	0,06	0,07	-0,12	0,00	0,01	0,03	0,08	0,21	2,19
ROE	0,19	0,59	-27,87	-0,03	0,03	0,10	0,25	0,73	30,04
ALV	1,06	1,90	-0,91	0,02	0,18	0,47	1,13	4,15	38,10
LIQ	2,11	2,58	0,00	0,48	1,04	1,41	2,22	5,82	75,55
Z	2,50	1,61	-2,41	0,24	1,41	2,27	3,33	5,38	29,72
RENT	0,12	0,10	0,00	0,02	0,06	0,10	0,16	0,32	2,72
TEI	0,24	0,24	-1,98	0,00	0,09	0,23	0,34	0,65	1,86
CFLOW	10,67	1,90	1,61	7,84	9,33	10,44	11,89	14,19	17,89

n = 86.158

De forma a examinar o impacto dos fatores determinantes nos indicadores da maturidade da dívida recorreu-se à correlação de Pearson. Na

Tabela 10 constata-se que não existe nenhuma relação forte entre as variáveis independentes e cada uma das variáveis dependentes, sendo que a maior é entre a DCP e a ALV ($r(86.156) = -0,49$). Destaca-se ainda, à semelhança dos estudos de Urbano (2011) e Proença (2012), a existência de uma relação negativa entre a LIQ e a DPC. As empresas que apresentam maiores níveis de liquidação tendem a não optar por dívida de curto prazo.

Percorrendo ainda a Tabela 10 verifica-se ainda que, e à semelhança de Hall et al. (2000) e Proença (2012), existe uma relação negativa ($r(86.156) = -0,27$) entre a EST e a DCP, ou seja, as empresas que apresentam um reduzido peso de ativos fixos tendem a optar pelo curto prazo. Por outro lado, a relação é positiva com a DLP ($r(86.156) = 0,33$).

Tabela 10 – Correlação de Pearson entre os indicadores da maturidade e os fatores determinantes

	EST	OC	DIMA	DIMVN	ROA	ROE	ALV	LIQ	Z	RENT	TEI	CFLOW
DCP	-0,266*	-0,092*	-0,082*	0,022*	-0,126*	0,049*	-0,488*	-0,269*	-0,269*	-0,114*	0,018*	-0,175*
DLP	0,331*	0,152*	0,040*	-0,041*	-0,144*	-0,004**	-0,239*	0,059*	-0,260*	-0,043*	-0,028*	0,003**

* $p < 0,05$; ** $p < 0,10$

A Tabela 11 compara a média dos indicadores da maturidade da dívida entre os tipos de empresa tendo em conta os critérios da UE. Verifica-se que a DCP vai decrescendo à medida que as empresas aumentam de dimensão, ou seja, as micro entidades apresentam uma DCP média de 49,9%, as pequenas de 49,5% e, por fim, as de média dimensão apresentam valores médios de 47%. Em termos da DLP verifica-se um comportamento semelhante ao da DCP, as micro entidades apresentam um valor de 16%, decrescendo ligeiramente para os 15,2% nas pequenas empresas e aumentando, nas empresas de média dimensão (15,5%). Em suma, a média da maturidade da dívida é diferente entre tipos de empresas (Proença, 2012).

Tabela 11 – Média dos Indicadores da Maturidade de acordo com a Dimensão das Empresas

	Micro	Pequena	Média	F (2, 86157)	p
DCP	0,50	0,50	0,47	44,54	0,001
DLP	0,16	0,15	0,16	4,10	0,017

A Tabela 12 apresenta os valores médios dos indicadores da maturidade tendo em conta o setor de atividade. A DCP oscila entre os 26% nas empresas pertencentes ao setor da eletricidade, gás, vapor, água quente e fria e os 55% no setor da construção. Em termos de DLP, os valores estão compreendidos entre os 14% dos setores da construção e das atividades de informação e comunicações e o valor de 32% no setor da eletricidade, gás, vapor, água quente e fria. Os valores encontrados são semelhantes aos de Proença (2012) e estão em linha com o estudo de Esperança et al. (2003) que afirma que, a média da maturidade da dívida é diferente entre os setores de atividade.

Tabela 12 – Média dos Indicadores da Maturidade de acordo com o setor de atividade

Setor	A	B	C	D	E	F	G	H	I	J	L	M	N	O	P	Q	R	S	F – Test (17, 86140)	P
DCP	0,42	0,39	0,48	0,26	0,44	0,55	0,53	0,5	0,45	0,51	0,28	0,46	0,53	0,27	0,5	0,37	0,47	0,44	148,52	0,001
DLP	0,21	0,19	0,16	0,32	0,22	0,14	0,14	0,2	0,22	0,14	0,28	0,15	0,15	0,26	0,2	0,18	0,18	0,19	86,169	0,001

4.3. Modelação

A opção por determinado algoritmo de DM não se apresenta como uma tarefa fácil. O utilizador deve ter em conta as características gerais que definem cada algoritmo. Cada algoritmo apresenta as suas vantagens e inconvenientes, distinguindo-se pelo método de otimização e estrutura de representação (Michell, 1997).

Na atualidade existe uma panóplia de técnicas de DM. No presente ponto são abordados os algoritmos: Regressão linear múltipla (RLM); Árvores de Decisão (AD) e Redes Neurais Artificiais (RNA), os mais utilizados em problemas de regressão e, por isso, selecionados para explicação da maturidade da dívida.

4.3.1. Regressão Linear Múltipla

A RLM apresenta-se como um método estatístico, utilizado principalmente em tarefas de previsão, que executa a análise de relações entre uma variável dependente (de saída) e as variáveis independentes (de entrada) (Pestana & Gageiro, 2005). O objetivo destes modelos pode ser explicativo (demonstrar uma relação matemática que pode indicar, não provando, uma causa-efeito) ou preditivo (obtenção de uma relação que permita perante futuras observações das variáveis independentes prever o valor da variável dependente sem necessidade de a medir) (Larose, 2005; Maimon & Rokach, 2010). Nesta técnica os dados são modelados tentando aproximá-los a uma linha reta, no caso da regressão simples e de um hiperplano no caso da RLM (Santos & Ramos, 2009).

A RLM é representada pela seguinte equação:

$$Y_i = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_n x_{ni} + \varepsilon_i \quad (2.1)$$

Y representa a variável dependente calculada a partir das variáveis independentes, $x_{1..n}$, sendo α e β os coeficientes da regressão. Estes coeficientes são apurados pelo Método dos Mínimos Quadrados (MMQ), método que procura minimizar a soma dos quadrados dos desvios (ε_i) existente entre os dados reais e os estimados pela equação (Santos & Ramos, 2009).

Pressupostos

Um dos principais pressupostos é a de que os erros do modelo têm de apresentar média nula, variância constante (homoscedasticidade) e não podem estar correlacionados entre si. Outra condição adicional é de que os erros têm de estar normalmente distribuídos. Os resíduos, devem distribuir-se aleatoriamente em torno do zero, tanto no modelo como para cada variável, caso não se verifique este pressuposto é necessário proceder a alterações no modelo, quer pode ser feito através da inclusão, eliminação ou transformação de variáveis. Por fim, é necessário ter em conta os problemas de multicolinearidade, ou seja, as variáveis devem ser linearmente independentes (Pestana & Gageiro, 2005).

Vantagens

A principal vantagem da regressão linear múltipla é a simplicidade da técnica, resulta do facto de ser exposta através de uma equação matemática que expressa a relação entre a variável dependente e as variáveis preditoras.

Desvantagens

Uma das principais desvantagens prende-se com a linearidade dos dados, ou seja, este tipo de regressão não consegue ser válida para dados não lineares ao contrário das AD e das RNA que se adequam a problemas com variadas dimensões e de fácil compreensão.

Parametrização do Modelo de Regressão Linear Múltipla

De forma a compreender que fatores mais contribuem para a explicação da maturidade da dívida nas PME de Portugal, Espanha, França, Itália, Grécia e Alemanha, utilizou-se em primeiro lugar a RLM. O estudo da RLM foi efetuado com recurso ao *software* IBM SPSS Modeler. Este programa faz a verificação e validação dos pressupostos enunciados anteriormente para a aceitação do modelo obtido, excluindo, desde logo, a possibilidade de existência de multicolinearidade entre as variáveis independentes. No presente estudo optou-se pela utilização do Método *Stepwise* para a construção dos modelos de regressão na medida em que este método exclui as variáveis que não satisfazem os critérios de seleção.

Os testes estatísticos estão desenhados de modo a controlarem o erro do tipo I, através da introdução de uma margem de erro (nível de significância - α), que se consubstancia na probabilidade de rejeição da hipótese nula sendo a mesma verdadeira. O método selecionado (*Stepwise*) utiliza o teste de significância no sentido de incluir nos modelos apenas as variáveis que apresentam poder explicativo. De acordo com o modelo *Stepwise* (Tabela 13) foram obtidos dois modelos de regressão: o primeiro modelo testa a relação entre a DCP e os fatores determinantes, no segundo modelo estudou-se a relação dos fatores determinantes mas com a DLP.

Os modelos integrantes da Tabela 13 integram apenas os fatores determinantes sem se ter em conta as variáveis de controlo sendo as equações do modelo representadas, respetivamente, por:

$$D\hat{C}P = 0,626 - 0,265 ESTi - 0,045 Zi - 0,008 LIQi - 0,027 ALVi - 0,052 DIMAi + 0,07 NDTSi - 0,048 CFLOWi + 0,094 DIMVNi + 0,165RENTi + 0,012ROEi - 0,064 ROAi + 0,003 TEIi$$

$$D\hat{L}P = 0,096 + 0,248 ESTi + 0,006 Zi + 0,07 LIQi - 0,031 ALVi + 0,048 DIMAi + 0,146 NDTSi - 0,011 CFLOWi - 0,039 DIMVNi + 0,30 RENTi + 0,003ROEi - 0,421 ROAi - 0,006 TEIi$$

Tabela 13 – Resultados da Regressão – Modelo Global

Variáveis Independentes	DCP				DLP			
	B	SE B	beta	t	B	SE B	beta	t
Constante	0,626	0,008		81,780	0,096	0,006		15,794
EST	-0,265	0,004	-0,221	-70,633*	0,248	0,003	0,289	82,667*
Z	-0,045	0,001	-0,281	-56,403*	0,006	0,001	0,055	9,859*
LIQ	-0,008	0,000	-0,148	-52,170*	0,007	0,000	0,175	55,212*
ALV	-0,027	0,001	-0,214	-51,337*	-0,031	0,000	-0,346	-74,383*
DIMA	-0,052	0,002	-0,366	-30,952*	0,048	0,001	0,474	35,904*
OC	0,070	0,034	0,010	2,050**	0,146	0,027	0,028	5,324**
CFLOW	-0,048	0,001	-0,351	-36,895*	-0,011	0,001	-0,113	-10,673*
DIMV	0,094	0,001	0,661	72,213*	-0,039	0,001	-0,038	-38,022*
RENT	0,165	0,019	0,064	8,523*	0,300	0,015	0,164	19,467*
ROE	0,012	0,001	0,046	16,944*	0,003	0,001	0,019	6,083*
ROA	-0,064	0,027	-0,018	-2,419**	-0,421	0,021	-0,169	-19,881**
TEI	0,003	0,001	0,008	2,836*	-0,006	0,001	-0,020	-6,726*
R ²	0,395				0,235			
R ² Ajustado	0,384				0,231			
F (12, 86146)	4471,81*				2159,19*			

* p < 0,01; **p < 0,05

Tendo em conta o coeficiente R² ajustado, os modelos expostos na Tabela 13, explicam 38% (F(12, 86146 = 4471,81, p < 0,01) da variabilidade da DCP em torno da média e 23% (F(12, 86146 = 2159,19, p < 0,01) da DLP. Os resultados obtidos para a DCP e DLP vão ao encontro dos resultados obtidos por Hall et al. (2000). Os valores do coeficiente de determinação pouco elevado são semelhantes aos resultados obtidos em estudos anteriores (por exemplo, Marques & Morgado, 2004; Proença, 2012 e Vieira &

Novo, 2010), apresentam tendo significância estatística ($p < 0,05$). Os pressupostos para a validação dos modelos foram cumpridos.

Em primeiro lugar, o teste do ANOVA representado pela estatística de teste F de *Snedecor* apresenta como hipótese nula que nenhum dos coeficientes seja significativo e, conseqüentemente, que nenhuma das variáveis independentes a eles associados não apresentarão poder explicativo com exceção para o termo independente (constante). Se os modelos forem globalmente significativos a evidência vai no sentido da rejeição da hipótese em teste. A significância é testada pelas seguintes hipóteses:

H0: $\beta_1 = \beta_2 = \beta_3 = \dots = \beta_{12} = 0$; (nenhuma das variáveis independentes contida no modelo explica a variabilidade da variável dependente).

H1: $i: \beta_i \neq 0$. ($i = 1, 2, 3, \dots, 12$ em que β_i é diferente de zero).

Para qualquer nível de significância (p) é rejeitada a hipótese nula pelo que, em suma, não existe evidência estatística de que os regressores não permitam descrever de forma apropriada a relação linear presente nos dados. Também o teste associado a cada coeficiente de regressão leva a rejeitar a hipótese nula de que o coeficiente é zero, pelo que se conclui que todos os fatores permitem explicar a maturidade.

Verificação das Hipóteses de Investigação

De acordo com os coeficientes obtidos (Betas) no modelo de regressão linear (Tabela 13) é possível observar a relação entre as variáveis relacionadas com a maturidade da dívida e os fatores determinantes. A Tabela 14 mostra o impacto expectável confrontado com o impacto observado dos fatores determinantes para explicar a maturidade da dívida de acordo com a informação retirada dos modelos de regressão linear múltipla.

A variável dimensão apresenta resultados antagónicos, ou seja, a dimensão medida pelo logaritmo do ativo ($B = - 0,05$ e $B = 0,05$) para a DCP e DLP respetivamente, valida a hipótese H1.a e H1.b. No entanto, se a medição for efetuada pelo logaritmo do volume de negócios ($B = 0,094$ e $B = - 0,039$ para a DCP e DLP) os resultados obtidos vão no sentido da rejeição das hipóteses H1.a e H1.b. Os resultados obtidos são semelhantes aos encontrados por Proença (2012). Por outro lado, contradizem os encontrados por Hall et al. (2000); Michaelas et al. (1999) e Vieira & Novo (2010). Estes resultados vão ao encontro do preconizado por Warner (1977), ou seja, as empresas de maior dimensão apresentam custos de financiamento relativamente inferiores aos das PME, facilitando

desta forma o acesso das grandes entidades a financiamento externo. A relação negativa entre a DIMA e a DCP revela a dificuldade que as PME enfrentam quando tentam aceder a financiamentos com maturidades mais elevadas o que faz com optem por financiamentos de curta duração.

Os resultados obtidos entre a EST e a maturidade da dívida vêm ao encontro das expectativas, ou seja, apontam para uma relação negativa com a DCP e positiva com a DLP. As empresas que apresentam níveis elevados de ativos fixos tangíveis tendem a recorrer a DLP. Estes resultados podem ser interpretados à luz da teoria do *Trade-off* e suportados pelos estudos de Scott (1977) e Titman & Wessels (1988). Os autores arguem que as entidades que apresentam mais ativos que possam servir de garantia em caso de falência tendem a utilizar mais dívida de longo prazo. A relação negativa com a DCP é explicável na medida em que as entidades com maior proporção de ativos tangíveis recorrem menos ao curto prazo. Outra das explicações para esta relação negativa pode ser o facto de as empresas utilizarem os seus ativos correntes que podem ser transformados facilmente em dinheiro para a liquidação de financiamentos de curta duração. Em suma, os resultados obtidos permitem a validação das hipóteses H2.a e H2.b tal como defendido por Hall et al. (2000); Vieira & Novo (2010) e Proença (2012).

As OC validam as hipóteses H3.a e H3.b, ou seja, as OC apresentam uma relação positiva quer com a DCP ($B = 0,07$) quer com a DLP ($B = 0,14$). Estes resultados validam os estudos de Michaelas et al. (1999), Mira & Gracia (2003) e Proença (2012). O resultado obtido contradiz a literatura existente (Barnea et al., 1980; Myers, 1977) porque as empresas que apresentam maiores níveis de gastos não desembolsáveis tendem a recorrer menos a dívida de longa duração. Estes resultados dão suporte à argumentação de Ross (1977) pois as empresas que apresentam maiores oportunidades de crescimento apresentam um menor risco pelo que conseguem obter condições de financiamento mais favoráveis.

A rendibilidade apresenta uma relação negativa com a DCP confirmando a teoria do *Pecking Order* preconizada por Myers (1984). Os resultados apontam para a validação das hipóteses H4.a e H4.b quer a variável utilizada seja a rendibilidade do ativo quer seja medida pelo resultado líquido (ROA) ($B = - 0,064$ e $B = - 0,421$ para a DCP e DLP). Caso se opte pela rendibilidade do capital próprio (ROE) e pela rendibilidade operacional (RENT) a relação é positiva ($B = 0,01$ e $B = 0,003$ para a DCP e DLP).

respetivamente) e ($B = 0,165$ e $B = 0,300$ para a DCP e DLP para a RENT). Estes valores são também defendidos na literatura (Hall et al., 2000; Esperança et al., 2003; Mira & Gracia, 2003 e Proença, 2012). Para além desta situação, os resultados obtidos sugerem ainda que as empresas preferem mais financiar-se no curto prazo do que no longo prazo na medida em que o efeito no longo prazo é maior do que no curto. Esta situação já havia sido defendida por Michaelas et al. (1999 e Proença (2012).

O risco apresenta, tal como esperado, uma relação negativa com a DCP, $B = -0,01$ (LIQ) e $B = -0,045$ (Z) e positiva com a DLP, $\beta = 0,007$ (LIQ) e $B = 0,006$ (Z) levando à validação das hipóteses H5.a e H5.b. Estes resultados confirmam que as empresas com maior risco tendem a recorrer mais a financiamentos com maturidades mais longas tal como evidenciado na literatura (Costa et al., 2014; Proença, 2012; Urbano, 2011). A relação negativa com a DCP é explicada pelo facto de as PME com maiores níveis de risco não conseguirem aceder a maturidades mais longas pelo que recorrem ao curto prazo para fazer face às suas obrigações (Proença, 2012).

O CFLOW revela a existência de uma relação negativa com a DCP e DLP, $B = -0,05$ (DCP) e $B = -0,01$ (DLP). Estes resultados para além de validarem as hipóteses H6.a e H6.b estão em linha com Mira & Gracia (2003) e Proença (2012). De acordo com a teoria do *Pecking Order*, as empresas tendem a financiar-se em primeiro lugar recorrendo aos fundos gerados internamente assim. Assim, a relação com a maturidade da dívida é negativa tal como obtido neste estudo. Estes resultados, no entanto, contradizem a teoria da agência, ou seja, se o *cash flow* livre aumentar, as maturidades decrescem, situação precisamente inversa ao preconizado na teoria anterior. Estes resultados vão ao encontro com o defendido por Poza & Kishida (2004), ou seja, a teoria da agência é irrelevante quando se trata do estudo das PME.

Por fim, TEI apresenta uma relação positiva ($B = 0,003$) com a DCP e negativa ($B = -0,006$) com a DLP. Estas conclusões não permitem validar a hipótese H.7. Apesar dos resultados encontrados a sua significância é relativamente reduzida quando comparada com as restantes variáveis pelo que a taxa de imposto apresenta pouco poder explicativo para a maturidade da dívida.

Tabela 14 – Impacto dos fatores determinantes na maturidade da dívida

	DCP		DLP	
	Expetativa	Observado	Expetativa	Observado
Dimensão	-	-/+	+	-/+
Estrutura dos Ativos	-	-	+	+
Oportunidades de Crescimento	+	+	+	+
Rendibilidade	-	-/+	-	-/+
Risco	-	-	+	+
Cash Flow	-	-	-	-
Taxa Efetiva de Imposto	+	+	+	-

4.3.2. Árvores de Decisão

As árvores de decisão tornaram-se populares devido à sua capacidade de exploração, identificação e classificação de estruturas complexas. As árvores de decisão são dos métodos mais adequados para problemas com muitas variáveis, na medida em que se trata de um método de previsão célere e não linear, e com uma clara representação e visualização do conhecimento. Este algoritmo tem como princípio subjacente o dividir para conquistar (Gama, Carvalho, Faceli, Lorena, & Oliveira, 2012). O modelo é composto por estruturas em árvore que representam conjuntos de decisões.

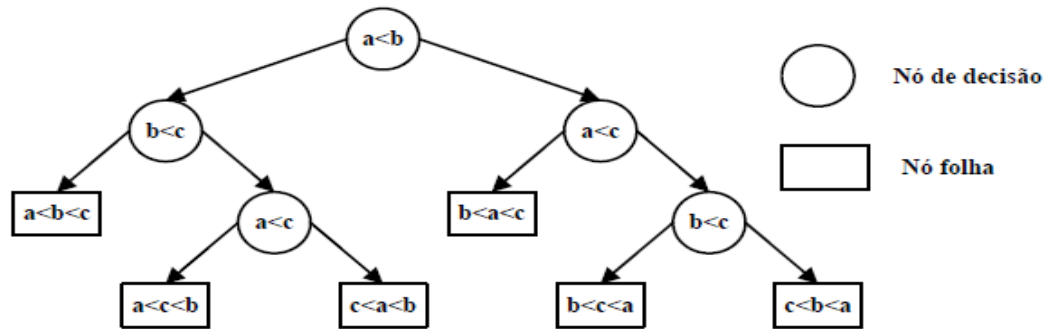
Existem vários algoritmos que arquitetam árvores de decisão destacando-se o AID – *Automatic Interaction Detection* (Morgan & Sonquist, 1963); CHAID – *CHi-squared Automatic Interaction Detection* (Hartingan, 1975; Kass, 1980); ID3 (Quinlan, 1979); o CART (Breiman, Friedman, Olshen, & Stone, 1984) e o C4.5 e C5 (Quinlan, 1988). Os dois primeiros modelos adaptam-se a problemas de classificação (variável alvo nominal) e regressão (variável alvo numérica), enquanto o último é apenas destinado a problemas de classificação (Gama et al., 2012). Dos modelos anteriores, o que mais se distingue é o CART – *Classification and Regression Trees*, apresentado por Breiman et al. (1984). Este modelo constrói árvores binárias facilmente interpretáveis, indicando ainda a importância relativa de cada variável independente.

O algoritmo CART integra a aprendizagem supervisionada sendo aplicável a problemas de classificação (dados qualitativos) e de regressão (dados quantitativos). Os componentes fundamentais de uma AD são os nós e as regras para a sua divisão. O crescimento da árvore inicia-se com o nó raiz que inclui a totalidade dos casos. De seguida, o algoritmo verifica qual a variável independente que melhor divide o nó pai em dois nós filhos. A variável selecionada em cada nó visa a maximização da pureza

dos dois nós filhos de modo a que cada grupo seja o mais homogéneo possível, ou seja, reduz a variabilidade da variável dependente intra-nós e maximiza a variabilidade inter-nós. Os nós terminais determinam a classificação ou predição determinadas pelo modelo através do valor médio da variável dependente que inclui todas as observações incluídas no nó (problemas de regressão) ou a proporção de casos corretamente classificados (problemas de classificação) (Larose, 2005; Laureano, 2009; Maimon & Rokach, 2010).

A Figura 4 mostra-nos um exemplo e o funcionamento de uma árvore de decisão.

Figura 4 – Exemplo de uma AD



Fonte: Caetano (2013)

O algoritmo CART, desenhado por Breiman et al. (1984) utiliza o critério *least squared deviation* (LSD) para definir as regras de divisão para problemas de regressão e a medida de impuridade de Gini para problemas de classificação.

Na modelação de uma AD é necessário ter em conta alguns critérios, ou seja, é necessário definir qual a melhor ramificação dos nós bem como as regras de paragem de processo de repartição a partir dos nós, passando a ser considerado nó-folha. É necessário, ainda, determinar-se o método de cálculo de cada nó-folha assim como a medidas adequadas ao desempenho do modelo (Ferreira, 1999; Gama et al., 2012).

Depois de se efetuar a construção de uma AD é obtida uma estrutura que se ajusta aos dados, normalmente algo extensa. Desta forma, é necessário aperfeiçoar-se o modelo recorrendo à técnica da poda que consiste fundamentalmente na eliminação dos ramos. Han & Kamber (2001) referem que a poda assume um papel de destaque na medida em que permite produzir árvores de menor dimensão mas que apresentam maior potencial e precisão quando aplicadas perante casos desconhecidos, apresentando uma maior capacidade de generalização. A poda pode ser efetuada de duas formas: durante ou após

a aprendizagem, no entanto, a maioria dos algoritmos recorre à segunda hipótese, ou seja, após a aprendizagem (Ferreira, 1999; Gama et al., 2012).

Na estratégia de pré corte a árvore é aparada quando se verifica a satisfação de algum dos critérios, limitando assim o seu crescimento, passando o nó a folha. Na estratégia pós-corte a árvore é construída e no final os ramos são cortados (Gama et al., 2012). Esta última estratégia baseia-se em duas medidas, erro de *backed-up* e erro estático. O primeiro corresponde ao somatório das classificações incorretas de todas as subárvores do corrente nó, Caso seja, maior ou igual ao erro seguinte, o nó é trocado para folha. O erro estático apresenta-se como o número de classificações incorretas tendo em consideração que todos os dados que chegam ao nó são classificados tendo em conta a classe maioritária da distribuição de classes (Gama et al., 2012). Todos estes métodos apresentam um equilíbrio entre a dimensão da árvore e a estimativa da taxa de erro. A estimativa da taxa de erro, à semelhança das formas de corte também apresenta duas estratégias: os métodos que estimam o erro a partir dos dados provenientes do conjunto de treino (erro de re-substituição) e os métodos que apuram o erro a partir de um conjunto não utilizado na construção da árvore (erro de generalização) (Gama et al., 2012).

A validação dos modelos de AD podem ser efetuada através da criação de um conjunto de teste, conjunto distinto do conjunto de treino extraído aleatoriamente do conjunto global, usualmente com dimensão igual ou inferior ao do conjunto de treino. O modelo apurado no conjunto de treino é depois aplicado ao conjunto de teste no sentido de se verificar a sua aplicabilidade a dados desconhecidos, ou seja, verificar a sua capacidade de utilização em tarefas de previsão (Santos & Ramos, 2009).

Quando a dimensão da amostra é reduzida não é possível dividi-la em conjunto de treino e teste opta-se pela validação cruzada. A validação cruzada é feita através da divisão da amostra em n subconjuntos de iguais dimensões de preferência apresentando a mesma distribuição das classes. O treino é efetuado através de várias iterações sobre o conjunto dos dados, sendo que em cada uma das iterações são decretados diferentes conjuntos de teste e treino. O erro é calculado a partir da construção das n árvores, sendo os resultados obtidos reflexo da totalidade do conjunto de dados o que permite tirar proveito da totalidade dos dados da amostra. O método mais utilizado na validação

cruzada é o *10-fold* que consiste em dividir a amostra em 10 subconjuntos (Cabete & Cardoso, 2006).

Vantagens

As AD apresentam as seguintes vantagens (Bose & Mahapatra, 2001; Gama et al., 2012: 114–115):

Flexibilidade – As AD obtêm resultados perante problemas que misturam variáveis contínuas, nominais e ordinais.

Seleção de variáveis – O processo de construção apura as variáveis a utilizar no modelo. Esta escolha permite a produção de modelos robustos relativamente ao acréscimo de variáveis redundantes ou insignificantes

Interpretação – Problemas complexos podem ser subdivididos em sub-problemas mais simples. As decisões baseiam-se nos valores das variáveis utilizadas na descrição da problemática;

Eficiência – Estes algoritmos são construídos de cima para baixo (*top-down*) utilizando a estratégia de dividir para conquistar sem retrocessos (*backtracking*).

Robustez – Este método adapta-se facilmente a dados omissos, sendo ainda inalterável perante a transformação de variáveis, como por exemplo, a logaritmização.

A RLM apresenta pressupostos para a sua aplicação enquanto as AD, por seu lado, constroem modelos para qualquer função desde que o número de exemplos seja suficiente; modelam funções não lineares sem ser necessário recorrer à transformação de variáveis, mantendo a facilidade de interpretação. Por fim, é bastante eficiente perante a presença de extremos, *outliers* bem como de atributos redundantes e omissos (Cabete & Cardoso, 2006).

Desvantagens

Apesar de apresentarem inúmeras vantagens as AD também apresentam alguns inconvenientes. Os mais abordados na literatura são (Bose & Mahapatra, 2001; Gama et al., 2012: 115–116):

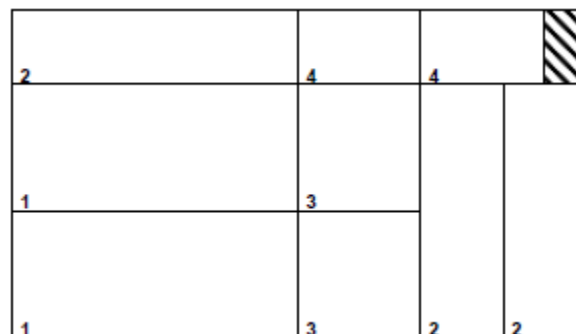
Replicação – Esta desvantagem refere-se ao facto de o algoritmo duplicar a «sequência de testes em diferentes ramos» de uma AD, conduzindo «a uma representação não concisa, que também tende a ter baixa precisão preditiva»;

Valores omissos – Se o valor da variável não for conhecido pode causar problemas na definição dos ramos a seguir. Desta forma, os algoritmos devem possuir mecanismos que tratem os omissos;

Instabilidade – Pequenas variações nos dados do conjunto de treino podem levar a grandes alterações ao nível da estrutura final da árvore. Em cada nó, o critério de divisão classifica as variáveis, sendo escolhido a melhor variável que divide os dados.

Ortogonalidade – O tipo de corte executado e o número de cortes permitidos são essenciais para problemas em que os objetos são retangulares ou em forma de blocos. Em problemas bidimensionais cujo padrão é definido através de cortes ortogonais a sua complexidade depende do número de mudanças nas direções de corte como também do número de corte paralelos. Na Figura 5 apresenta-se um padrão ortogonal bidimensional com três cortes, quatro variáveis e no canto superior direito aparece a perda.

Figura 5 – Exemplo da Ortogonalidade



Os dados que apresentam ruído levantam dois problemas, em primeiro lugar, a classificação a dados desconhecidos pode não ser confiável, as estatísticas apuradas nos nós mais profundos tendem a não ter grande importância devido ao reduzido número de valores classificados (Gama et al., 2012). No entanto, estes nós refletem de melhor forma o conjunto de treino (sobreajustamento) e aumentam o erro. O segundo problema surge devido à grande amplitude das árvores, o que as torna de difícil compreensão, assim sendo a poda ajuda a aparar os nós mais profundos minimizando significativamente estes problemas (Gama et al., 2012). Cabete & Cardoso (2006) referem ainda que se se estiver perante um caso de sobre ajustamento no crescimento da AD aos valores constantes do conjunto de treino, situação que dificulta a generalização do modelo construído, podem ser definidas regras de paragem do crescimento que permitem a obtenção de melhores resultados.

Parametrização do Modelo CART – Árvore de Decisão

Na aplicação da metodologia das AD procedeu-se à utilização do algoritmo CART já explicado precedentemente. Com base nas partições da amostra e nas variáveis evidenciadas na parte da preparação e compreensão dos dados e tendo como objetivo a construção de um modelo que apresente o menor erro possível, foram testados diferentes parâmetros, dos quais se optou pela seleção daquele que melhores resultados apresentou para cada uma das duas variáveis dependentes (DCP e DLP) e que vão ser analisados de seguida.

Maturidade de Curto Prazo

O modelo que melhores resultados apresentou é composto pela seguinte parametrização: a partição da amostra foi distribuída em 50% para o conjunto de treino e 50% para o conjunto de teste ⁵ (Chan, Fan, Prodromidis, & Stolfo, 1999); o número máximo de níveis de profundidade da árvore (10); mínimo de observações por nível no nó pai (1%); mínimo de observações por nível de nó filho (0,5%) e com poda (Anexo 5).

Do modelo gerado (Anexo 4) constatou-se a existência de uma profundidade de apenas oito níveis e um número de nós filho de trinta e três. Este modelo apresentou resultados bastante satisfatórios, em termos do R^2 (embora o R^2 seja mais utilizado na avaliação das regressões lineares, foi utilizado neste ponto para efetuar a comparação inicial entre os desvios observados na variável dependente nas restantes técnicas, sendo no capítulo seguinte efetuada a comparação entre as restantes medidas de avaliação), na ordem dos 83% no conjunto de treino e 80% no conjunto de teste. O modelo seguinte (Anexo 6) foi aquele que melhores resultados apresentou para a DCP.

Neste modelo realizou-se a poda (*prunning*). Esta técnica permite tornar a árvore mais genérica evitando desta forma os problemas de sobreajustamento. Se a árvore crescer até ao seu máximo, esta refletirá todos os detalhes dos dados, o que para além de a tornar mais complexa torna-a menos precisa em dados de teste, na medida em que poderá apresentar resultados específicos para um caso específico.

No modelo gerado constata-se que apenas as variáveis LIQ, ALV e EST integram a árvore de decisão. A árvore é decomposta em primeiro lugar pela variável LIQ. A

⁵ Os resultados obtidos foram ligeiramente superiores à partição da amostra em 70% para o conjunto de treino e 30% para o conjunto de teste (repartição *standard*).

amostra é dividida nas empresas que apresentam um rácio de liquidação (ativos correntes/passivos correntes) inferior ou igual (63% da amostra) e superior (37% da amostra) a 1,726. A DCP apresenta um valor previsto de 0,622 (nó 1) para as empresas que estão no limite inferior e de 0,276 (nó 2) no limite superior.

O nó 57 (Anexo 5) é que apresenta maior DCP prevista e o nó 14 apresenta a menor maturidade. O nó 57 integra 1,15% das PME com uma maturidade prevista de 1,767. Este nó apresenta as seguintes características: a LIQ é menor ou igual a 1,261; a ALV apresenta um valor negativo menor ou igual a 0,451 e a EST é inferior ou igual a 0,276. As empresas que integram este nó apresentam um rácio de alavancagem negativo o que pressupõe que as mesmas apresentam capitais próprios negativos. A relação entre os ativos e passivos correntes é menor ou igual a 1,767 pelo que as empresas conseguem solver as suas obrigações de curto prazo. Estas empresas apresentam dificuldades de financiamento de longo prazo porque para além dos capitais próprios negativos ainda apresentam um peso relativamente reduzido de ativos colateralizáveis.

Por outro lado, o nó 14 (Anexo 5) é onde se encontra a DCP prevista mais reduzida (0,106) integrando 7,06% das observações. Neste nó a única variável independente é a LIQ, ou seja, as empresas que exibem um valor de ativos correntes superior em 4,8 vezes ao passivo corrente tendem a apresentar a menor maturidade de curto prazo. Em suma, estas empresas conseguem solver facilmente as suas dívidas de curto prazo com recurso à transformação em fluxos de caixa dos seus ativos correntes.

Maturidade de Longo Prazo

No estudo da DLP a AD que melhores resultados apresenta inclui a seguinte parametrização (Anexo 6 **Erro! A origem da referência não foi encontrada.**), em termos de níveis optou-se por um número máximo de 10, em termos de número mínimo de observações nos nós pai e filho selecionaram-se 1% e 0,5% respetivamente. A poda esta foi também escolhida para integrar o modelo.

O modelo resultante desta parametrização (Anexo 7) apresenta uma profundidade de dez níveis e um número de nós filho de 32. A variabilidade da DLP em torno da média, calculada através do coeficiente de determinação R^2 , situa-se nos 0,62 quer no conjunto de treino quer para o conjunto de teste. As variáveis que integram o modelo para além das identificadas no modelo anterior, LIQ, ALV e EST, inclui ainda Z. A primeira

variável a dividir a amostra não é a LIQ, como no modelo anterior, mas sim a ALV que divide a amostra em empresas que apresentam um rácio de alavancagem menor ou igual (nó 1) e maior que 0,764 (nó 2).

O nó 14 (anexo 7) apresenta a menor (0,03) e o 51 (anexo 7) a maior (2,76) DLP. No nó 14 que inclui 9,26% das observações, a maturidade é explicada através de três variáveis, ou seja, as empresas apresentam valores iguais ou inferiores a 2,34 para a ALV e 2,88 para o Z. Estas empresas apresentam capitais próprios superiores aos passivos. Para além disto, situam-se na zona de conforto do risco, ou seja, as empresas apresentam um risco mínimo de falência. Estes resultados validam a teoria do *Pecking Order* porque as empresas financiam-se em primeiro lugar através do autofinanciamento e só depois recorrem á dívida. No outro extremo, aparece o nó 51 com uma DLP estimada de 2,76. Estas empresas apresentam capitais próprios negativos, tal como já havia sido referenciado para a DCP. Este nó inclui 0,5% das observações e é composto por empresas que apresentam uma ALV negativa (-0,721); uma EST menor ou igual a 0,384 e uma LIQ igual a 0. Estas empresas encontram-se em situação de falência porque para além de apresentarem dificuldades no que diz respeito aos capitais próprios não têm ativos correntes suficientes para fazer face às obrigações de curto prazo. Em suma, estas entidades não conseguem aceder a financiamentos de curto prazo tendo que recorrer a financiamentos de longo prazo dando como garantias os seus ativos tangíveis.

Em suma, o algoritmo CART mostra uma excelente capacidade de previsão tendo-se obtido um desempenho bastante satisfatório com uma elevada consistência e resultados facilmente interpretáveis, embora o R^2 da DLP seja bastante menor do que o conseguido com o DCP. Desta forma, o algoritmo CART poderia, por si só, ser um método adequado para a exploração do problema em estudo – maturidade da dívida. Nas AD não foram efetuados modelos para as variáveis de controlo porque as não se apresentam como estatisticamente significativas para o estudo da maturidade da dívida.

4.3.3. Redes Neurais Artificiais

As RNA surgiram em meados da década de 40 e início da década de 50 do século XX com os estudos de McCulloch & Pitts (1943). Os autores construíram um neurónio artificial. Estes neurónios executam funções lógicas simples sendo que cada uma pode executar diferentes funções. Estes sistemas produzem um elevado poder computacional na medida em que qualquer problema pode ser representada por uma combinação de

funções e é suscetível de ser representada por uma rede de neurónios (Gama et al., 2012).

A compreensão do funcionamento do cérebro humano permitirá que se usufrua das suas reais valências. Nos últimos 50 anos, cerca de 10% dos prémios Nobel atribuídos tiveram na sua base estudos relacionados com o funcionamento das sinapses, neurónios e de outras estruturas das RNA (Rojas, 1996). Atualmente, as RNA desempenham um papel de extrema importância no processamento de dados, sendo a sua principal diferença em relação à programação tradicional, o facto de permitir aprendizagem, fiabilidade e paralelismo e em consequência a generalização. A generalização é possível pelas aproximações com grau de confiabilidade, mesmo que os dados sejam distintos dos utilizados no conjunto de treino. Dado isto, as RNA são algoritmos bastante utilizados na realização de tarefas complexas nomeadamente no reconhecimento de padrões ou aproximações de funções (Rojas, 1996).

As RNA inspiram-se no cérebro humano que consistem em conjuntos intimamente ligados de neurónios. Embora um neurónio por si só apresente uma estrutura simples, o mesmo não acontece quando vários neurónios se juntam e criam uma rede densa e complexa capaz de executar tarefas complexas tais como classificação, ou reconhecimento de padrões. O cérebro humano é composto aproximadamente por 1.011 neurónios cada um ligado a 10.000 outros neurónios perfazendo cerca de 10^{15} conexões sinápticas (Larose, 2005). As RNA inspiram-se na estrutura e funcionamento do sistema nervoso, tendo como finalidade a simulação da capacidade de aprendizagem do cérebro humano para a obtenção de conhecimento, tentando a um nível elementar reproduzir o tipo de aprendizagem não-linear que surge nas redes de neurónios encontradas nos humanos (Gama et al., 2012).

Perceptrão

Rosenblatt (1958) desenvolveu o perceptrão, conjunto de neurónios artificiais e sinapses interligadas. Cada neurónio calcula o somatório da multiplicação entre os sinais que aparecem da camada de entrada e os pesos associados. As ponderações são armazenadas nas sinapses, sendo que o resultado da ponderação é utilizado como entrada de uma função de ativação (incilamente do tipo linear), e por sua vez, a saída consubstancia-se

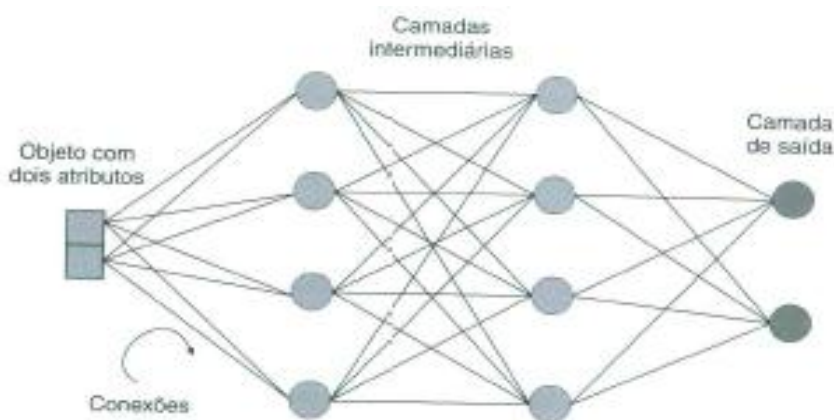
no resultado do neurónio. A saída poderá ser ou não alvo de processamento por outro neurónio (Rosenblatt, 1958).

Uma das limitações das redes que utilizam uma camada é que apenas conseguem ser exequíveis para problemas que classifiquem objetos linearmente separáveis. Existem diversas estruturas de RNA as quais diferem essencialmente pelo número de nós intermédios. As redes do tipo *perceptron* não apresentam qualquer nível intermédio, apenas possuem nós de entrada e saída, o que permite uma aprendizagem mais simples, no entanto, estas redes estão condicionadas a algumas tarefas, mais concretamente a tarefas aproximáveis a problemas lineares. As restantes redes, as que apresentam nós intermédios, designam-se por perceptrão multinível permitindo a sua utilização a tarefas não lineares (Santos & Ramos, 2009).

Estrutura

A rede neuronal mais utilizada baseia-se na arquitetura de múltiplas camadas, *Multi-Layer Perceptron* (MLP) (Papert & Minsky, 1969). Esta arquitetura é composta por três camadas essenciais (Figura 6): a primeira é a camada de entrada (*input layer*), esta camada inclui os dados seleccionados, na camada de processamento (*middle layer*) onde é feito o processamento dos dados e por fim, a camada de saída (*output layer*) é composta pelos resultados obtidos. A camada intermédia pode ser composta por várias camadas de neurónio ou por nenhuma, havendo assim, uma ligação entre a camada de entrada e a de saída.

Figura 6 – Redes Multicamadas

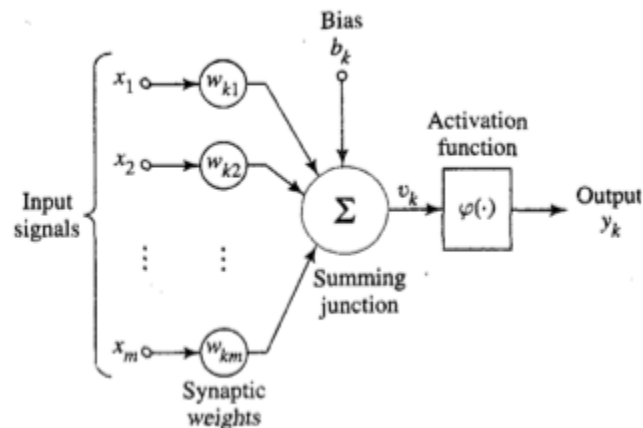


Fonte: Gama, Carvalho, Faceli, Lorena, & Oliveira (2012: 133)

Um neurónio consiste numa unidade de processamento essencial para as RNA podendo apresentar valores negativos e positivos, na medida em que constituem as suas unidades básicas. Este elemento apresenta três características básicas (Graupe, 2007; Haykin, 1999; Rojas, 1996): É constituído por um conjunto de sinapses que efetuam a ligação entre as diferentes unidades de processamento (neurónios). Cada um apresenta um peso associado, que é um fator para a regulação dos valores processados; Um somatório que possibilita fazer a operação sobre os valores de entrada e os pesos das respetivas sinapses, tendo estas operações o mesmo número de combinações lineares; Uma função de ativação que limita a amplitude de resultados do neurónio, permitindo que este apresente um valor máximo e mínimo uniformizando os resultados o que permite a sua comparação.

A estrutura dos neurónios pode ser visualizada na Figura 7. Na figura é feita referência ao *bias*. Este componente tem como objetivo aumentar ou reduzir a entrada da rede da função de ativação, dependendo se apresenta um valor negativo ou positivo (Graupe, 2007; Haykin, 1999; Rojas, 1996).

Figura 7 – Constituição do Neurónio



Fonte: Haykin (1999)

Camadas

A primeira camada, de entrada, é aquela que recebe os dados selecionados para serem processados (Graupe, 2007; Haykin, 1999). A camada intermédia faz o processamento dos dados tal como a camada de saída que processa os valores de forma a obter os resultados. A única camada que não processa os dados é a de entrada, nesta camada os

neurónios não interferem com os resultados obtidos. As camadas intermédias são também designadas de escondidas, porque não interagem com o ambiente em redor, interagem apenas com as camadas de neurónios da mesma rede (Graupe, 2007; Haykin, 1999). Graupe (2007) refere ainda que apesar de ser o mais usual, a camada intermédia não necessita de ter o mesmo numero de neurónios da camada antecedente.

A camada de saída é a ultima fase de procesamento a que os dados estão sujeitos. Tal como evidenciado para a camada anterior, a de saída não necessita apresentar o mesmo número de neurónios que as duas camadas anteriores, normalmente espera-se que apresente um único valor. O valor é dependente das funções utilizadas (Graupe, 2007; Haykin, 1999). As funções mais utilizadas são a linear; tangente hiperbólica, sigmoide e a logistica (Graupe, 2007; Haykin, 1999).

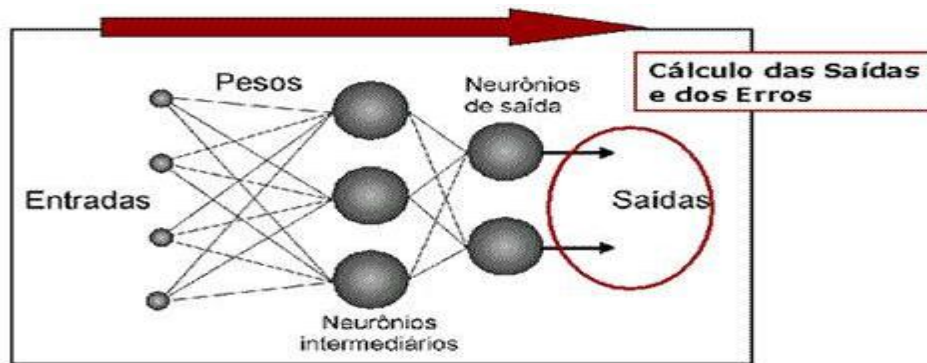
Papert & Minsky (1969) provaram matematicamente as limitações do perceptrão de Rosenblatt (1958) e conjecturaram que o *Multi-Layer Perceptron* sofria das mesmas limitações. Dadas as limitações dos modelos as RNA foram abandonadas durante a década de 70 do século passado. No entanto, o estudo das redes neuronais ressurgiu durante a década de 80 com a descoberta da aplicação do método de *backpropagation* de Rumelhart, Hinton, & Williams (1986) ao *Multi-Layer Perceptron*.

Algoritmo Backpropagation

Este algoritmo de aprendizagem de RNA foi proposto por Paul Werbos em 1970, tornando-se conhecido através dos estudos de Rumelhart e McClland em 1986. Estes modelos surgiram da necessidade existente de encontrar uma forma do modelo de MLP reduzir o erro na camada de saída. Rumelhart, Hinton, & Williams (1986) utilizaram a retro-propagação do erro como forma de minizar o erro na camada de saída, ajustando os pesos através de métodos, tal como o gradiente descendente.

O processo de aprendizagem do *backpropagation* subdivide-se em duas fases (Gama et al., 2012).: *Forward* – a partir dos pesos aleatoriamente gerados pelo investigador, para a ligação das diferentes variáveis preditivas no sentido entrada-saída da rede, o algoritmo considera estes pesos fixos e calcula, para cada observação, o valor do *output* estimado de acordo com a expressão do neurónio de saída. De seguida (Figura 8) este valor é comparado com o valor real e obtem-se assim o erro.

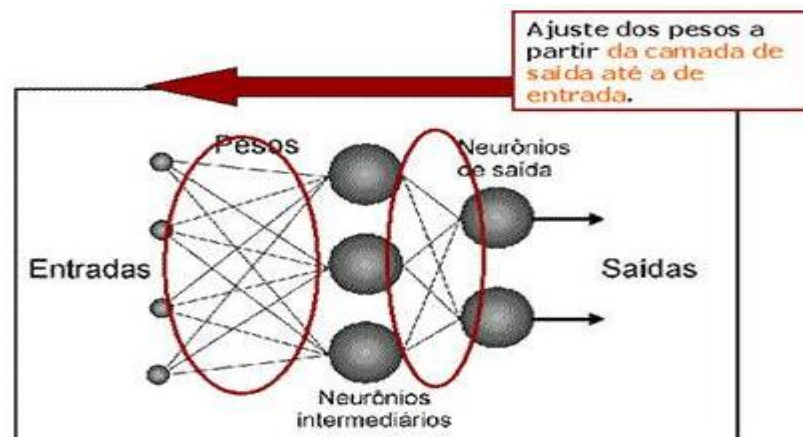
Figura 8 – Fase de Propagação



Fonte: Oliveira (2009: 40)

Backward – Com base no erro anteriormente calculado, o algoritmo procede à aprendizagem e regressa novamente ao início da rede tentando ajustar os pesos em cada ligação iniciando de seguida o processo (Figura 9). O ajustamento é feito pelo método do Gradiente Descendente, utilizando a informação de qual deve ser a direcção do ajustamento dos pesos de forma a minimizar o erro de cada neurónio.

Figura 9 – Fase da *Backpropagation*



Fonte: Oliveira (2009:40)

Os pesos são atualizados em cada observação até ser completado o ciclo, ou seja, depois de terminadas as fases identificadas anteriormente (*Forward* e *Backward*), à totalidade dos elementos presentes no conjunto de treino. Este número de ciclos poderá ser definido pelo investigador de modo a evitar problemas de sobreajustamento e limitar o tempo de computação (Gama et al., 2012; Oliveira, 2009b).

À medida que a aprendizagem vai avançando, a rede começa a ficar cada vez mais precisa no que diz respeito à sua utilização para dados desconhecidos. Estes passos são replicados até que o nível de desempenho do modelo seja aceitável. Depois de a rede estar treinada é testada no conjunto de teste de forma a validar o modelo em dados desconhecidos (Santos & Ramos, 2009).

Vantagens

As RNA apresentam características que as distinguem das outras técnicas e métodos (Cortez, 2002; Gama et al., 2012; Haykin, 1999): aprendizagem e generalização (conseguem uma descrição do todo a partir de partes, constituindo-se como uma forma eficiente de armazenamento e aprendizagem de conhecimento); processamento paralelo (tarefas complexas são realizadas num curto espaço de tempo); não linearidade (resolve problemas de natureza não linear); adaptabilidade (adapta-se de acordo com as mudanças do ambiente); robustez (capacidade de processamento de dados com ruído ou informação incompleta de forma eficiente e ainda conseguem manter o desempenho quando existe a desativação de alguns nodos) e flexibilidade (apresentam um grande domínio de aplicação).

Estas características fazem com que as RNA se apresentem como uma ferramenta alternativa no reconhecimento, classificação e previsão na área das finanças, principalmente devido aos níveis elevados de precisão, adaptabilidade, robustez eficácia e eficiência na resolução de problemas financeiros (Wong & Selvi, 1998).

Desvantagens

As RNA, ao contrário das AD que se baseiam em regras intuitivas compreensíveis para a maioria dos utilizadores da informação devido ao grafismo do *output* na medida em que cria fronteiras de decisão ortogonais que podem se facilmente traduzidas em regras de decisão (Gama et al., 2012), não são totalmente aceites (designados por caixas negras) devido à falta de explicação dos modelos gerados, nomeadamente quantas camadas apresenta ou quais as funções de ativação, e a sua relação com os dados fornecidos (Graupe, 2007; Haykin, 1999). Berry & Linoff (2004) referem que esta técnica apresenta dificuldades perante a existência de um elevado número de variáveis de entrada, ou seja, um elevado número de variáveis pode levar à demora no tempo de treino prejudicando os resultados do modelo.

Nas RNA, como em todos os modelos preditivos, é necessário ter em conta os problemas de sobreajustamento. As RNA são relativamente boas para dados de treino, no entanto, nem sempre o são para dados desconhecidos, sendo uma das grandes problemáticas das RNA uma vez que a interpretação do funcionamento do modelo é extremamente difícil. Uma das formas existentes para controlar os problemas de sobreajustamento é limitar o número de conexões uma vez que representam a complexidade dos modelos. Contudo, não existe uma técnica/teoria que permita identificar qual o número ideal de conexões. Uma das soluções pode passar pela divisão da amostra em conjunto de treino e teste, sendo o primeiro utilizado na concepção do modelo e o segundo na deteção da eficácia da predição (Gama et al., 2012).

Parametrização do Modelo de Redes Neuronais Artificiais

A modelação das RNA (Anexo 8) teve em conta as mesmas variáveis que integraram a RLM e as AD. Tal como nas AD, também nas RNA foram testados diversas alternativas tendo-se optado pela apresentação do modelo com os resultados mais relevantes para o estudo da maturidade da dívida. Em primeiro lugar optou-se pela partição da amostra de acordo com a literatura, ou seja, 70% para o conjunto de treino e 30% para o conjunto de teste. No entanto, os resultados obtidos foram mais relevantes quando se escolheu a partição da amostra em 50% para treino e teste (Chan et al., 1999).

O número de épocas definido para o conjunto de treino foi de mil, ainda que se tenham testado outras alternativas, os resultados obtidos foram inferiores, quer para o conjunto de treino quer para o de teste. O método de otimização utilizado foi o Gradiente Descendente na medida em que é aquele que produz melhores resultados e é o mais eficiente quando se está perante amostras de grandes dimensões como é a amostra do estudo em causa. A função de ativação utilizada na camada intermédia e de saída foi a Tangente Hiperbólica, esta função acelera a convergência do algoritmo de treino da rede neuronal (Pereira, 1999).

Foram testados diferentes modelos, quer com uma camada intermédia quer com duas, não se verificando diferenças relevantes no conjunto de teste porém, o modelo com resultados mais relevantes apresenta apenas uma camada intermédia. Na literatura não existe um critério que permita identificar qual o número de neurónios ideal na camada intermédia. De um modo geral, quanto menor for o número de neurónios melhor é o poder de generalização, reduzindo os problemas de sobreajustamento. Por outro lado, as

redes que apresentem poucos neurónios podem não ter capacidade de modelação e aprendizagem dos dados em problemas complexos, ocorrendo problemas de subajustamento, ou seja, a rede não converge durante o ciclo de treino (Pereira, 1999). De acordo com a informação anterior optou-se pela não limitação das unidades em cada camada. A taxa de aprendizagem foi fixada nos 0,1, na medida em que é o valor mais recomendado pela literatura, esta taxa permite incrementos mais reduzidos em cada ciclo (Pereira, 1999). Apesar disto, testaram-se modelos com uma taxa de aprendizagem de 0,2, todavia os resultados obtidos foram menos consistentes.

As redes que melhores resultados apresentaram para cada uma das variáveis são compostas pelos parâmetros incluídos na Anexo 6 com os resultados constantes da Tabela 16.

Discussão dos Resultados das RNA

Em termos dos neurónios, para a DPC, o modelo apresenta 8 neurónios e na DLP foram identificados 10 neurónios. Em termos de importância das variáveis (Tabela 15) conclui-se que para a DCP é explicada em cerca de 50% por apenas quatro variáveis (ALV, LIQ, ROA e DIMA). Na DLP, por sua vez, a ALV e LIQ também apresentam um peso significativo (20%), no entanto, a terceira variável mais importante é a EST comprovando a importância que os ativos fixos apresentam junto da DLP. Em suma, conclui-se que as variáveis mais importantes relacionadas com a maturidade da dívida são a ALV, seguida da LIQ que mede o risco, da rentabilidade do ativo e do capital próprio (ROA e ROE), da EST bem como da TEI e da dimensão das empresas. Tal como nas AD a localização, o tipo e o setor de atividade das empresas não se apresentam como variáveis importantes para o estudo da maturidade da dívida pelo que não se efetuou a análise das RNA tendo em conta as variáveis de controlo.

Tabela 15 – Importância das variáveis

	ALV	LIQ	ROA	DIMA	DIMV N	TEI	EST	RENT	Z	OC	ROE	SETA T	CFLO W	TIPO	TRAB	PAIS
DCP	14%	13%	13%	10%	7%	7%	6%	6%	5%	5%	4%	3%	3%	2%	1%	1%
DLP	20%	20%	4%	7%	7%	5%	10%	5%	6%	3%	6%	3%	3%	0%	0%	0%

Em termos de resultados (Tabela 16) é possível observar através do resultado do coeficiente de determinação R^2 que a variância da DCP é explicada em 81% no conjunto de treino decrescendo ligeiramente para os 78% em teste. A DLP, tal como verificado, nos modelos anteriores apresenta uma relação menor na ordem dos 70% quer para treino quer para teste.

Tabela 16 – Resultados das Redes Neurais

		DCP	DLP
Ajustamento Conjunto de Treino	R^2	81%	70%
Ajustamento Conjunto de Teste	R^2	78%	70%

Tal como verificado para as AD, também as RNA se apresentam bastante adequadas para a resolução da maturidade da dívida, dado que o desempenho obtido é bastante satisfatório.

Capítulo 5 – Avaliação e Resultados

5.1. Avaliação

Esta etapa divide-se em dois passos – em primeiro lugar realiza-se a avaliação dos modelos e em segundo lugar procede-se à revisão de todo o processo de DM de forma a verificar-se se foram excluídos ou omitidos fatores importantes. No final do subcapítulo é tomada a decisão sobre qual o melhor modelo a utilizar-se no estudo da maturidade da dívida.

Na literatura existem diferentes medidas que podem ser utilizadas na avaliação dos modelos de DM. Neste ponto são apenas abordados algumas das métricas de avaliação existentes para problemas de regressão dado ser o tipo de problema que se pretende estudar. Witten et al. (2011) referem que as principais medidas de avaliação para problemas de regressão são: o erro absoluto médio (EAM), o erro quadrático médio (EQM); a raiz quadrada do erro quadrático médio (RQEQM) e o coeficiente de determinação R^2 .

O erro absoluto médio efetua a média da magnitude dos erros individuais sem ter em conta o sinal dos mesmos. Este método apresenta como principal relevância o facto de ser relativamente pouco influenciado pela presença de *outliers* nos dados, isto resulta do facto de todos os erros serem examinados proporcionalmente à sua magnitude. O cálculo do EAM é feito através da seguinte fórmula (5.1). Os valores previstos são comparados com os valores inicialmente observados.

$$EAM = \frac{|p_1 - a_1| + \dots + |p_n - a_n|}{n} \quad (5.1)$$

O erro quadrático médio é muito usado nas técnicas de regressão, sendo, ao contrário do anterior, sensível à presença de *outliers*. O cálculo baseia-se na seguinte fórmula (5.2).

$$EQM = \frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n} \quad (5.2)$$

A raiz quadrada do erro quadrático médio corresponde à raiz quadrada do EQM, sendo expressa pela seguinte equação (5.3).

$$RQEQM = \sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}} \quad (5.3)$$

Os modelos de previsão são mais precisos, ou seja, apresentam melhores resultados, quanto mais reduzidos forem os valores do EQM e do RQEQM.

O coeficiente de determinação R^2 apresenta-se como um bom indicador para a comparação de modelos de regressão. O coeficiente mostra a distribuição dos valores previstos e observados através de uma linha ou curva. O valor do R^2 oscila entre o 1 e o $-\infty$ e quanto mais próximo de 1 menor é a diferença entre os valores observados e previstos pelo que melhor é o modelo produzido.

Apresentação da Avaliação dos Modelos

Na explicação da maturidade da dívida é possível concluir que o risco, a rendibilidade, a dimensão, a estrutura dos ativos e a alavancagem das PME apresentam um papel relevante na explicação da maturidade da dívida.

Na Tabela 17 estão representadas as métricas dos modelos que apresentaram os melhores resultados no que diz respeito à DCP e DLP. Em termos do coeficiente de determinação R^2 , as AD apresentam-se para a DCP como o algoritmo onde se encontraram os melhores resultados com uma correlação de 0,83 em treino e de 0,80 em teste. Em termos de erros as AD e as RNA apresentam erros semelhantes. Na RLM apesar da reduzida correlação entre as variáveis já era possível observar que a ALV, a EST, o CFLOW, as variáveis relacionadas com o risco (LIQ e Z) bem como as relacionadas com a dimensão (DIMA e DIMV) apresentavam um peso significativo na explicação da maturidade da dívida. No entanto, o modelo de AD apenas inclui as variáveis LIQ, ALV e EST. Nas RNA as variáveis mais importantes são a ALV, LIQ, ROA, DIMVN e DIMA. Dos resultados obtidos conclui-se que a tomada de decisão por parte dos gestores é influenciada nos financiamentos de curto prazo pelo risco, pela alavancagem bem como pela dimensão das empresas. A estrutura também influencia a tomada de decisão na medida em que os ativos fixos têm a capacidade de servir de garantia na obtenção de financiamentos.

A DLP, à semelhança da DCP, é explicada em parte pela ALV, LIQ, Z, EST e DIMA. Dos modelos selecionados para o estudo da DLP é possível observar pela Tabela 17 que o modelo que melhores resultados produziu foi a RNA com um R^2 de 70% em treino e em teste. As AD também apresentam um melhor coeficiente de determinação na ordem dos 62% contra os 24% da RLM. Em termos de EMQ e da RQEQM as diferenças são pouco significativas entre as RNA e as AD, existindo uma diferença de aproximadamente 2% na RQEQM. As variáveis que integram o modelo de AD são a ALV, LIQ, Z e EST. As RNA para além destas refere ainda que a dimensão é também significativa para a explicação da DLP. Em suma, a tomada de decisão por parte dos gestores ao nível da escolha da DLP é explicada em parte pelas variáveis que compreendem o risco das empresas (LIQ e Z), bem como pela alavancagem. A dimensão das empresas quer seja medida pelo volume de negócios quer seja medida pelo ativo é também importante para a tomada de decisão tal como o peso dos ativos colateralizáveis na estrutura de ativos das PME. Apesar dos resultados obtidos pelas RNA serem ligeiramente superiores optou-se pela escolha do modelo de AD como o que melhor explica a DLP. Esta escolha prende-se fundamentalmente com a facilidade de visualização e interpretação dos resultados dada a dificuldade/opacidade dos resultados provenientes das RNA.

Tabela 17 – Métricas Obtidas para a Maturidade de Curto Prazo

		DCP			DLP		
Modelo		RLM	AD	RNA	RLM	AD	RNA
Ajustamento Conjunto de Treino	EAM	0,16	0,07	0,08	0,12	0,08	0,07
	EQM	0,04	0,01	0,01	0,03	0,01	0,01
	RQEQM	0,20	0,11	0,11	0,16	0,12	0,10
	R^2	0,40	0,83	0,81	0,24	0,62	0,70
Ajustamento Conjunto de Teste	EAM	0,16	0,07	0,08	0,12	0,08	0,07
	EQM	0,04	0,01	0,02	0,03	0,01	0,01
	RQEQM	0,21	0,12	0,12	0,16	0,12	0,10
	R^2	0,37	0,80	0,78	0,23	0,61	0,70

5.2. Implementação

Este trabalho de cariz académico torna-se relevante na medida em que permite identificar quais os fatores que mais contribuem para a tomada de decisão da maturidade da dívida nas PME. No entanto, devido à complexidade do problema alvo

de estudo e do limite temporal não é possível fazer a implementação do estudo a casos concretos.

5.3. Discussão e Resultados

A revisão da literatura sobre a maturidade da dívida nas PME permitiu constatar a não existência de estudos anteriores onde tivessem sido utilizadas técnicas de DM tais como as AD, mais concretamente o algoritmo CART, e as RNA, o que invalida a comparação dos resultados obtidos com base nas métricas utilizadas. A generalização dos fatores determinantes da maturidade da dívida é bastante complexa na medida em que apenas foi analisada uma pequena porção das PME existentes nestes países. A avaliação sobre a proporção da variância explicada pelos modelos não era expectável apresentar valores elevados dado que os estudos existentes utilizam maioritariamente a RLM com resultados semelhantes aos obtidos neste trabalho.

Baseado numa amostra aleatória composta por 86.158 PME relativas ao ano de 2013 e recorrendo à metodologia CRISP-DM e a técnicas de DM – RLM, AD e RNA, os resultados apontam para que os fatores que mais influenciam a tomada de decisão dos gestores nas PME são para a DCP e para a DLP, a alavancagem, estrutura dos ativos e o risco. De entre os fatores analisados apenas as oportunidades de crescimento e impostos apresentam resultados fracos na escolha da maturidade da dívida. Os resultados apontam também para que as empresas optem em primeiro lugar pelo financiamento interno verificando-se a Teoria do *Pecking Order*. Por outro lado, a Teoria da Agência não se verifica nas PME, tal como defendido por Poza & Kishida (2004) na medida em que na maioria dos casos o empresário é também o gestor.

De acordo com o primeiro objetivo evidenciam-se as seguintes conclusões:

Em termos de RLM observou-se que a relação entre a EST é negativa com a DCP e positiva com a DLP. Estes resultados verificam a Teoria do *Trade-off* que sugere que as empresas com maiores níveis de ativos tangíveis tendem a apresentar maiores níveis de dívida, isto porque estes ativos podem ser utilizados como garantia em caso de falência das empresas. Também nas AD e RNA a estrutura de ativos é importante na explicação da maturidade da dívida sendo de destacar que nas RNA a importância aumenta do curto prazo para o longo prazo concluindo-se que a estrutura dos ativos influencia a tomada de decisão dos gestores.

Os resultados relativos à dimensão indicam à semelhança da estrutura dos ativos, relações opostas, negativa e positiva com a DCP e DLP respetivamente. Esta situação ocorre se a variável utilizada for o ativo. Se se utilizar o volume de negócios para medir a dimensão de uma empresa os resultados obtidos são inversos, ou seja, uma relação positiva com a DCP e negativa com a DLP. As empresas que apresentam um maior volume de ativos tendem a financiar-se no longo prazo dado que conseguem menores custos de financiamento. Por outro lado, as empresas com maiores volumes de negócios tendem a optar pelo curto prazo devido à sua maior capacidade operacional. A importância da dimensão é também verificada nas RNA sendo mais relevante para a DCP. Contudo, a dimensão não integra os modelos de AD pelo que não é importante neste caso para a explicação da maturidade.

Em termos do risco, obteve-se uma relação positiva para a DLP e negativa para a DCP. As PME que apresentam maior risco tendem a financiar-se no longo prazo. Por outro lado, as empresas que apresentam dificuldades no cumprimento das obrigações de curto prazo, em consequência dos problemas escassez de fundo maneio, recorrem ao curto prazo de forma a cumprirem as suas obrigações. As técnicas de DM evidenciam também a importância destas variáveis principalmente no que diz respeito ao longo prazo.

As OC apresentam uma relação positiva com qualquer que seja a maturidade da dívida. As empresas com maiores oportunidades de crescimento anunciam junto dos credores a sua qualidade pelo que não enfrentam problemas de *financial distress* o que lhes permite a obtenção de condições de crédito mais favoráveis. No entanto, esta variável apresenta uma importância relativamente reduzida no que concerne às técnicas de DM pelo que não se apresenta como importante para a explicação da maturidade da dívida.

A rendibilidade apresenta uma relação negativa com a DCP e DLP se medida pela rendibilidade do ativo. Estes resultados sugerem que as PME financiam em primeiro lugar os seus projetos com recurso aos resultados gerados e só depois recorrem a dívida devido aos custos financeiros inerentes à segunda opção. Posto isto, verifica-se a teoria do *Pecking Order*. Se a variável for a rendibilidade do capital próprio os resultados apontam para uma relação positiva na medida em que à medida que as empresas aumentam o volume de dívida o custo do capital próprio diminui aumentando a rendibilidade para os proprietários. À semelhança das oportunidades de crescimento,

também a rendibilidade não é um fator determinante para a explicação da maturidade da dívida.

A alavancagem é o fator que mais contribui para a explicação da maturidade da dívida, quer se esteja perante a DCP quer seja a DLP. A tomada de decisão por parte dos gestores nas PME tem em conta a alavancagem, ou seja, o peso que a dívida e o capital próprio apresentam junto das origens de capitais. A alavancagem reduz-se à medida que aumenta o volume de dívida pelo que a relação terá de ser negativa com a DCP e DLP. Este peso é mais evidente quando se trata da DLP de acordo com os resultados obtidos nas AD e RNA.

O *Cash flow* apresenta uma relação negativa com a DCP e DLP. Esta situação permite a validação da teoria do *Pecking Order* tal como descrito para a rendibilidade. No entanto, e à semelhança dos impostos, a variável não apresenta uma influência significativa, quer nas AD quer nas RNA, para a explicação da maturidade da dívida nas PME.

As variáveis de controlo, setor de atividade, país e tipo de empresas não se apresentaram como significativas nos modelos de DM. Assim, conclui-se que apesar de a medida da maturidade da dívida variar de setor para setor, de país para país e por tipo de empresa, não existe evidência estatística para que se possa concluir que estas variáveis sejam determinantes para a tomada de decisão nas PME.

Capítulo 6 – Conclusão

Neste último capítulo, são expostas as principais conclusões derivadas deste trabalho e a relevância do estudo para as PME europeias, para a comunidade académica bem como para o público em geral. São ainda apresentadas as principais limitações evidenciadas ao longo do estudo bem como sugestões para futuras investigações.

6.1. Principais Conclusões

Este estudo empírico propôs-se a investigar que fatores afetam a tomada de decisão relativa à maturidade da dívida nas PME não financeiras de Portugal, Espanha, França, Itália, Alemanha e Grécia. Posto isto, identificou-se que a estrutura dos ativos, a alavancagem, o cash flow, dimensão, a rendibilidade, as oportunidades de crescimento e o risco são os fatores que a literatura sugere como aqueles que o gestor mais tem em conta aquando da tomada de decisão. Outros dos objetivos era verificar se os fatores afetam de igual forma os diferentes países e se varia de setor para setor de atividade e de dimensão para dimensão.

Em suma, este trabalho permitiu evidenciar os principais fatores que os gestores das PME têm em conta aquando da tomada de decisão relativamente à escolha da maturidade da dívida. Os fatores estão relacionados fundamentalmente com a estrutura dos ativos, alavancagem e risco. Em termos das teorias referidas na revisão da literatura é possível observar que as teoria do Trade-off e do *Pecking Order* apresentam um papel importante na definição da maturidade da dívida. Por outro lado, os resultados não apontam para a validação da teoria da agência nas PME.

Para o segundo objetivo conclui-se que apesar da média, variar de setor para setor, de país para país e por tipos de empresa, não existem evidências estatísticas para que estes fatores sejam tidos em conta pelo gestor na tomada de decisão nas PME.

6.2. Importância e Contributo do Estudo

A estrutura de capitais e mais concretamente a maturidade da dívida das empresas tem constituído, nas últimas décadas, um dos temas de maior interesse na literatura financeira. Desde o estudo de Modigliani & Miller (1958), vários estudos e teorias têm

sido desenvolvidas nesta área, inicialmente aplicadas nas grandes empresas, e entretanto também aplicado ao universo das PME.

Neste estudo pretendeu-se preencher algumas das lacunas existentes na literatura, seguindo-se as indicações e limitações de alguns estudos. Em primeiro lugar, optou-se por uma amostra que contempla informação financeira de empresas pertencentes a diferentes categorias (micro, pequenas e médias) e de diferentes países (Alemanha, Portugal, Espanha, Itália, Grécia e França). As empresas selecionadas cobrem a totalidade dos setores de atividade, com exceção dos setores financeiros, que dadas as especificidades existentes nestes setores foram excluídas da amostra.

Este estudo representa um importante contributo para a literatura financeira, para a comunidade académica e para o público em geral na medida em que foca um dos principais problemas com os gestores das empresas se deparam atualmente, ou seja, o excesso de endividamento. Neste estudo foi feita uma abordagem profunda sobre a temática utilizando diferentes técnicas estatísticas com recurso a técnicas de DM para além da técnica tradicional de regressão linear múltipla.

O estudo apresenta-se como um dos pioneiros no que diz respeito à utilização de técnicas de DM para a explicação da maturidade da dívida na literatura financeira.

6.3. Principais Limitações e Investigações Futuras

Este estudo, tal como todos os outros, apresenta limitações que poderão ser aproveitadas para futuras investigações. Em primeiro lugar, no estudo são consideradas as principais variáveis explicativas da maturidade da dívida encontradas na literatura. No entanto, podem existir outras variáveis que apresentem um papel importante para a explicação da maturidade da dívida. Por limitação temporal e computacional, não foi possível avaliar o impacto de fatores internos organizacionais e as características dos países das PME na definição da maturidade da dívida bem como de uma amostra de maiores dimensões.

A literatura é ainda escassa sobre a temática da maturidade da dívida nas PME. Em futuros estudos seria pertinente alargar a comparação não só entre países da Europa mas também pertencentes a outros continentes e avaliar se os fatores determinantes são semelhantes aos da Europa.

Por fim, dada a utilidade das técnicas de DM na descoberta de novo conhecimento no campo da previsão da maturidade da dívida sugere-se a utilização de outras técnicas tais como as Máquinas de Vetores de Suporte (Vapnik, 1995).

Referências Bibliográficas

- Agrawal, R., & Srikant, R. 1994. Fast Algorithms for Mining Association Rules. *Intl. Conf. on Very Large Databases*: 487–499.
- Alcarva, P. 2011. *O guia completo sobre A BANCA E AS PME*. (Vida Económica, Ed.) (1ª Edição.). Porto.
- Amadeus. 2015. <https://amadeus.bvdinfo.com/version-2015821/home.serv?product=amadeusneo>.
- Ang, J. 1991. Small business uniqueness and the theory of financial management. *The journal of small business finance*, 1(1): 1–13.
- Antoniou, A., Guney, Y., & Paudyal, K. 2006. The Determinants of Debt Maturity Structure: Evidence from France, Germany and the UK. *European Financial Management*, 12(2): 161–194.
- Barnea, A., Haugen, R. A., & Senbet, L. W. 1980. A Rationale for Debt Maturity Structure and Call Provisions in the Agency Theoretic Framework. *The Journal of Finance*, XXXV(5): 1223–1235.
- Berry, M., & Linoff, G. 2004. *Data Mining Techniques - For Marketing, Sales, and Customer Relationship Management* (2nd Editio.). Indiana: Wiley Publishing, Inc.
- Bhaird, C., & Lucey, B. 2010. Determinants of capital structure in Irish SMEs. *Small Business Economics*, 35(3): 357–375.
- Bose, I., & Mahapatra, R. K. 2001. Business data mining — a machine learning perspective. *Information & Management*, 39(3): 211–225.
- Brealey, R., Myers, S., & Allen, F. 2007. *Princípios de Finanças Empresariais* (8ª Edição.). Madrid: McGraw-Hill.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. . 1984. *Classification and Regression Trees*. (Wadsworth, Ed.). Belmont, CA.
- Brick, I. E., & Ravid, S. A. 1985. On the Relevance of Debt Maturity Structure. *The Journal of Finance*, XL(5): 1423–1438.
- Cabete, N. P., & Cardoso, M. G. M. S. 2006. Algoritmo CART : Previsão do Desempenho na Matemática do Secundário. *Revista de Ciências da Computação*, I: 11–27.
- Caetano, N. 2013. *Previsão de tempos de internamento de pacientes via técnicas de Data Mining*. ISCTE Business School.

- Chan, P. K., Fan, W., Prodromidis, A., & Stolfo, S. 1999. Distributed Data Mining in Credit Card Fraud Detection. *IEEE Intelligent System's Special Issue on Data Miningsue on Data Mining*, 1–17.
- Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinart, T., Shearer, C., et al. 2000. *CRISP-DM 1.0 - Step-by-step data mining guide*. (SPSS, Ed.).
- Cios, K. J., Pedrycz, W., Swiniarski, R. W., & Kurgan, L. 2007. The Knowledge Discovery Process. In Hardcover (Ed.), *Data Mining - A Knowledge Discovery Approach*: 9–22.
- Comissão Europeia. 2003. *Recomendação n° 2003/361/CE de 20 de Maio de 2003*.
- Comissão Europeia. 2007. *Decreto-Lei n° 372/2007 de 6 de Novembro de 2007*.
- Comissão Europeia. 2009. *Decreto-Lei n° 143/2009 de 16 de Junho de 2009*.
- Correia, S. 2012. *As Determinantes da Maturidade da Dívida*. Faculdade de Economia - Universidade do Porto.
- Correia, S., Brito, P., & Brandão, E. 2014. *Corporate Debt Maturity An International Comparison of Firm Debt Maturity Choices*. No. 544.
- Cortez, P. 2002. *Previsão de Séries Temporais. PhD - UMinho*. Universidade do Minho - Escola de Engenharia.
- Costa, S., Laureano, L. M. S., & Laureano, R. M. S. 2014. The Debt Maturity of Portuguese SMEs: The Aftermath of the 2008 Financial Crisis. *Procedia - Social and Behavioral Sciences*, 150: 172–181.
- Couto, G., & Ferreira, S. 2009. *Os Determinantes da Estrutura de Capital das Empresas do PSI-20*. No. CEEApIA WP 08/2009.
- Damodaran, A. 2002. *Investment Valuation: Tools and Techniques for Determining the Value of Any Asset*. John Wiley & Sons.
- Daskalakis, N., & Thanou, E. 2010. An analysis of the capital structure decisions of micro, small and medium firms: The case of Greece. *SSRN*, 1683161.
- De Miguel, A., & Pindado, J. 2001. Determinants of capital structure: new evidence from Spanish panel data. *Journal of Corporate Finance*, 7(1): 77–99.
- DeAngelo, H., & Masulis, R. W. 1980. Optimal capital structure under corporate and personal taxation. *Journal of Financial Economics*, 8(1): 3–29.
- Demirgüç-Kunt, A., & Maksimovic, V. 1998. Law, Finance, and Firm Growth. *The Journal of Finance*, LIII(6): 2107–2137.

- Desai, V. S., Crook, J. N., & Overstreet, G. a. 1996. A comparison of neural networks and linear scoring models in the credit union environment. *European Journal of Operational Research*, 95(1): 24–37.
- Diamond, D. W. 1991. Debt maturity structure and liquidity risk. *The Quarterly Journal of Economics*, 106(3): 709–737.
- Diamond, D. W. 1993. Seniority and maturity of debt contracts. *Journal of Financial Economics*, 33(3): 341–368.
- Dias, C. 2012. *Estrutura de Capital e Determinantes da Rendibilidade das Empresa do Distrito de Santarém*. Instituto Politécnico de Tomar - Escola Superior de Gestão de Tomar.
- Drobtz, W., & Fix, R. 2005. What Are the Determinants of the Capital Structure ? Evidence from Switzerland. *Swiss Journal of Economics and Statistics*, 141(I): 71–113.
- Durand, D. 1952. Costs of Debt and Equity funds for business: Trends and Problems of measurement. *National Bureau of Economic Research*: 215–262.
- Esperança, J. P., Gama, A. P. M., & Gulamhussen, M. A. 2003. Corporate debt policy of small firms: an empirical (re)examination. *Journal of Small Business and Enterprise Development*, 10(1): 62–80.
- Fama, E. F., & French, K. R. 2002. Testing Trade-Off and Pecking Order Predictions about Dividends and Debt.pdf. *The Review of Financial Studies*, 15(1): 1–33.
- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. 1996. From data mining to knowledge discovery in databases. *AI magazine*, 37–54..
- Ferreira, M. 1999. Árvores de Regressão e generalizações - Aplicações - Árvores de Regressão e generalizações - Aplicações -. *Tese mestrado em Matemática Aplicada - Universidade do Porto*.
- Gama, A. P. 2000. *Os Determinantes da Estrutura de Capital das PME's industriais portuguesas*. Porto: Associação da Bolsa de Derivados do Porto.
- Gama, J., Carvalho, A., Faceli, K., Lorena, A., & Oliveira, M. 2012. *Extracção de Conhecimento de Dados - Data Mining*. (E. Sílabo, Ed.). Lisboa.
- García-Teruel, P. J., & Martínez-Solano, P. 2010. Ownership structure and debt maturity: new evidence from Spain. *Review of Quantitative Finance and Accounting*, 35(4): 473–491.
- Graupe, D. 2007. *Principal of Artificial Neural Netwroks*. (Word Scientific, Ed.) (2nd ed.). Chicago.
- Guedes, J., & Opler, T. I. M. 1996. The Determinants of the Maturity of Corporate Debt Issues. *The Journal of Finance*, LI(1): 1809–1834.

- Hall, G., Hutchinson, P., & Michaelas, N. 2000. Industry effects on the determinants of unquoted SME's capital structure. *International Journal of the Economics of Business*, 7(3): 297–312.
- Han, J., & Kamber, M. 2001. *Data Mining: Concepts and Techniques*. (M. K. Publishers, Ed.).
- Harris, M., & Raviv, A. 1990. Capital Structure and the Informational Role of Debt. *The Journal of Finance*, XLV(2): 321–349.
- Harris, M., & Raviv, A. 1991. The Theory of Capital Structure. *The Journal of Finance*, XLV(1): 297–355.
- Hartigan, J. A. 1975. *Clustering Algorithms*. Nova Iorque: John Wiley & Sons.
- Hastie, T., Tibshirani, R., & Friedman, J. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Elements* (2nd ed.).
- Haykin, S. 1999. *Neural Networks - A Comprehensive Foundation*. (P. Hall, Ed.) (2nd Editio.).
- Hernández-Cánovas, G., & Koëter-Kant, J. 2005. *SME Financing in Europe : Cross-Country Determinants of Debt Maturity SME Financing in Europe : Cross-Country Determinants of Debt Maturity*. VU University Amsterdam, Faculty of Economics, Business Administration and Econometrics.
- Holmes, S., & Kent, P. 1991. An Empirical Analysis of the Financial Structure of Small and Large Australian Manufacturing Enterprises An Empirical Analysis of the Financial. *The Journal of Entrepreneurial Finance*, 1(2): 141–154.
- Hovakimian, A., Opier, T., & Titman, S. 2001. The Debt-Equity Choice. *Journal of Financial & Quantitative Analysis*, 36(March): 1–25.
- Jensen, M. C. 1986. Agency Cost Of Free Cash Flow, Corporate Finance, and Takeovers. *American Economic Review*, 76(2).
- Jensen, M. C., & Meckling, W. H. 1976. Theory of the Firm : Managerial Behavior , Agency Costs and Ownership Structure Theory of the Firm : Managerial Behavior , Agency Costs and Ownership Structure. *Journal of Financial Economics*, 3(4): 305–360.
- Johnson, S. 2003. Debt Maturity and the Effects of Growth Opportunities and Liquidity Risk on Leverage. *Review of Financial Studies*, 16(1): 209–236..
- Kass, G. V. 1980. An Exploratory Technique for Investigating Large Quantities of Categorical Data. *Applied Statistics*, 29(2): 119.
- Kim, E. H. 1974. A Mean-variance theory of optimal capital structure and corporate debt capacity. *Journal of Finance*, 33(1962): 45–63.

- Kohonen, T. 1982. Self-organized formation of topologically correct feature maps. *Biological Cybernetics*, 43(1): 59–69.
- Kraus, A., & Litzenberger, R. H. 1973. A state-preference model of optimal financial leverage. *Journal of Finance*, 28(4): 911–922.
- Larose, D. T. 2005. *Discovering Knowledge in Data: An Introduction to DATA MINING*. New Jersey: John Wiley & Sons, Inc. Wiley.
- Laureano, L. 2009. *Essays on debt maturity*. ISCTE Business School.
- Leland, H. E., & Toft, K. B. 1996. Optimal Capital Structure, Endogenous Bankruptcy, and the Term Structure of Credit Spreads. *Journal of Finance*, 51(3): 987–1020.
- Lev, B. 1969. Industry Averages as Targets for Financial Ratios. *Journal of Accounting Research*, 7: 290–299.
- Lewis, C. 1990. A multiperiod theory of corporate financial policy under taxation. *Journal of Financial and Quantitative Analysis*, 25(1): 25–44.
- López-Gracia, J., & Mestre-Barberá, R. 2011. Tax effect on Spanish SME optimum debt maturity structure. *Journal of Business Research*, 64(6): 649–655.
- Maimon, O., & Rokach, L. 2010. *Data Mining and Knowledge Discovery Handbook* (2nd Editio.). London: Springer.
- Marques, F., & Morgado, A. 2004. A Maturidade da dívida das Pequenas e Médias empresas industriais portuguesas. *X Congresso de Contabilidade*. Estoril.
- Matias, F., & Esperança, J. P. 2009. *Finanças Empresariais*. (Texto Editores, Ed.).
- McCulloch, W. S., & Pitts, W. 1943. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4): 115–133.
- Meggison, W. L. 1997. *Corporate Finance Theory*. Reading, Massachusetts.: Addison Wesley.
- Méndez, V. G. 2013. Determinants of debt maturity structure across firm size. *Revista Española de Financiación y Contabilidad*, XLII(158): 187–209.
- Michaelas, N. F., Chittenden, F., & Poutziouris, P. 1999. Financial Policy and Capital Structure Choice in UK SMEs: Empirical Evidence from Company Panel Data. *Small Business Economics*, 12: 113–130.
- Michell, T. 1997. *Machine Learning*. USA: McGraw Hill.
- Mira, F., & Gracia, J. 2003. *Pecking Order versus Trade-off: An empirical approach to the small and medium enterprise capital structure*.

- Modigliani, F., & Miller, M. H. 1958. The Cost of Capital Corporate Finance and Theory of Investment. *The American economic*, 48(3): 261–297.
- Modigliani, F., & Miller, M. H. 1963. Corporate Income Taxes and the Cost of Capital: A Correction. *The American Economic Review*, 53(3): 433–443.
- Morgan, J. N., & Sonquist, J. a. 1963. Problems in the Analysis of Survey Data, and a Proposal. *Journal of the American Statistical Association*, 58(302): 415–434.
- Morris, J. 1976. On corporate debt maturity strategies. *The Journal of Finance*, XXXI(I): 29–38.
- Morris, J. R. 1992. *Factors affecting the maturity structure of corporate deb*. Denver.
- Mota, A. G., Barroso, C. D., Nunes, J. P., & Ferreira, M. A. 2012. *Finanças da Empresa - Teoria e Prática*. (M. Robalo, Ed.) (Edições Sí.). Lisboa.
- Muller, P., Gagliardi, D., Caliandro, C., Bohn, N. U., & Klitou, D. 2014. *Annual Report on European SMEs 2013/2014 – A Partial and Fragile Recovery*.
- Myers, C. S. 1977. Determinants of Corporate Borrowing. *Journal of Financial Economics*, 5(2): 147–175.
- Myers, S. C. 1984. The Capital Structure Puzzle. *The Journal of Finance*, 39(3): 575–592.
- Myers, S. C. 2001. Capital Structure. *Journal of Economic Perspectives*, 15(2): 81–102.
- Myers, S. C. 2002. Financing of Corporations. *Handbook of the Economics of Finance*, 1–41.
- Myers, S. C., & Majluf, N. S. 1984. *Corporate financing and investment decisions when firms have*. No. 1396.
- Oliveira, J. P. A. 2009. *Identificação E Caracterização De Situações De “Churn” em Sistemas De Telecomunicações*. ISCTE - IUL.
- Ozkan, A. 2000. An empirical analysis of corporate debt maturity structure. *European Financial Management*, 6(2): 197–212.
- Ozkan, A. 2002. The determinants of corporate debt maturity: evidence from UK firms. *Applied Financial Economics*, 12(1): 19–24.
- Palma, A. R. da. 2011. *Determinantes da Maturidade da dívida: Evidência empírica de Portugal e Espanha*. Universidade Técnica de Lisboa - Instituto Superior de Economia e Gestão.
- Papert, S., & Minsky, M. 1969. *Perceptrons: An Introduction to Computational Geometry*. The MIT Press.

- Pereira, B. B. 1999. *Introduction to Neural Networks in Statistics: Center of Multivariate Analysis*. Penn.
- Pereira, J. A., Rebelo, E., & Coelho, L. 2004. Determinantes da estrutura de capital: Uma abordagem ao caso das empresas pertencentes aos países do mercado Euronext. *Faculdade de Economia da Universidade do Algarve*, 48.
- Pestana, M., & Gageiro, J. 2005. *Análise de Dados para Ciências Sociais: A Complementaridade do SPSS*. (E. Sílabo, Ed.) (4^a Edição.). Lisboa.
- Pontes, S., & Laureano, R. 2012. Proposta de Teoria de suporte ao relato financeiro diferenciado. *Proceedings of the XV Encuentro AECA 'Nuevos caminos para Europa: El papel de las empresas y los gobiernos: 1–22*.
- Portelinha, M. 2005. *Redes Neurais Artificiais para a Previsão da Qualidade em Carnes*. Universidade do Minho.
- Poza, E., & Kishida, R. 2004. Does the family business interaction factor represent a resource or a cost? *Family Business Review*, 17(2): 99–118.
- Proença, P. 2012. *Determinants of Capital Structure and Financial Crisis impact: Evidence from portuguese SMEs*. ISCTE Business School - Instituto Universtário de Lisboa.
- Quinlan, J. R. 1979. Discovering rules by induction from large collections of examples. *Expert Systems in the Microelectronic Age*, 168–201.
- Quinlan, R. 1988. *C4.5: Programs for Machine Learning*. Morgan Kaufmann Publishers, Inc.
- Rajan, R. G., & Zingales, L. 1995. What Do We Know about Capital Structure ? Some Evidence from International Data. *The Journal of Finance*, 50(5): 1421–1460.
- Rasiah, D., & Kim, P. K. 2011. A Theoretical Review on the use of the Static Trade off Theory, the Pecking Order Theory and the Agency Cost Theory of Capital Structure. *International Research Journal of Finance and Economic*, 63: 150–159.
- Rojas, R. 1996. *Neural Networks - A Systematic Introduction*. (Springer, Ed.)*Neural Networks*.
- Rosenblatt, F. 1958. The perceptron: A probabilistic model for information storage and organization in *Psychological Review*, 65(6): 386–408.
- Ross, S. A. 1977. The determination of financial structure : the incentive-signalling approach. *The Bell Journal of Economics*, 8(1): 23–40.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986. Learning internal representations by error propagation. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*.

- Santos, M. F., & Azevedo, C. S. 2005. *Data mining : descoberta de conhecimento em bases de dados* (1ª Edição.). Lisboa: FCA - Editora de Informática, Lda.
- Santos, M. Y., & Ramos, I. 2009. *Business Intelligence - Tecnologias da Informação na Gestão de Conhecimento* (2ª edição.). FCA - Editora Informática.
- Sarkar, S. 1999. Illiquidity Risk, Project Characteristics, and the Optimal Maturity of Corporate Debt. *The Journal of Financial Research*, XXII(3): 353–370.
- Scott, J. H. 1977. Bankruptcy, Secured Debt, and optimal capital structure. *Journal of Finance*, XXXII(1): 1–19.
- Silva, A. M. G. 2007. *O desempenho escolar via uma abordagem de descoberta de conhecimento em bases de dados*, Universidade do Minho.
- Stohs, M., & Mauer, D. 1996. The Determinants of Corporate Debt Maturity Structure. *Journal of Business*, 69(1): 279–312.
- Stulz, R. 1990. Managerial discretion and optimal financing policies. *Journal of Financial Economics*, 26(1): 3–27.
- Tan, P.-N., Steinbach, M., & Kumar, V. 2006. *Introduction to Data Mining*. (Addison-Wesley, Ed.).
- Titman, S., & Wessels, R. 1988. The determinants of capital structure choice. *The Journal of Finance*, 43(1): 1–20.
- Urbano, H. 2011. *Debt Maturity Structure across europe: Evidence from Greece, Ireland, Italy, Portugal and Spain*. ISCTE Business School - Instituto Universitário de Lisboa.
- Vapnik, V. 1995. *Statistical Learning Theory*. (Springer-Verlag, Ed.), Nova Iorque.
- Vieira, E. S., & Novo, A. J. 2010. A Estrutura de Capital das PME: evidência no mercado português. *Estudos do ISCA*, 2(2): 1–19.
- Warner, J. 1977. Bankruptcy costs: Some Evidence. *The Journal of Finance*, 32(2): 337–348.
- Weston, J. F. 1989. What MM Have Wrought. *The Journal of the Financial Management Association*, 18(2): 29–38.
- Witten, I. H., Frank, E., & Hall, M. a. 2011. *Data Mining: Practical Machine Learning Tools and Techniques. Complementary literature None* (2nd ed.)..
- Wong, B. K., & Selvi, Y. 1998. Neural network applications in finance: A review and analysis of literature (1990–1996). *Information & Management*, 34(3): 129–139.
- Xu, G., Li, L., & Zhang, Y. 2011. *Web Mining and Social Networking*. (Springer, Ed.).

Anexos

Anexo 1 – Resumo dos Atributos	86
Anexo 2 – Descrição das Variáveis	87
Anexo 3 – Setores de Atividade	88
Anexo 4 – Parâmetros da Árvore de Decisão - Maturidade de Curto Prazo	89
Anexo 5 – Árvore de Decisão – Maturidade de Curto Prazo	90
Anexo 6 – Parâmetros da Árvore de Decisão - Maturidade de Longo Prazo	91
Anexo 7 – Árvore de Decisão – Maturidade de Longo Prazo	92
Anexo 8 – Parâmetros das Redes Neurais	93

Anexo 1 – Resumo dos Atributos

Atributo	Tipo
Ativo Total	Contínua
Ativos Fixos tangíveis	Contínua
Ativo Corrente	Contínua
Passivo Corrente	Contínua
Passivo Não Corrente	Contínua
Passivo Total	Contínua
Capital Próprio	Contínua
Volume de Negócios	Contínua
Resultado Antes de Juro, Amortizações, Depreciações e Impostos	Contínua
Depreciações	Contínua
Juros	Contínua
Resultado Antes de Imposto	Contínua
Resultado Líquido	Contínua
Resultados Transitados	Contínua
Cash Flow	Contínua
Número de Colaboradores	Contínua
Localização das PME	Nominal
Tipo de Empresas – Categorias PME	Nominal
Setor de Atividade	Nominal

Anexo 2 – Descrição das Variáveis

Variável	Descrição	Tipo
DCP	Endividamento de Curto Prazo	Contínua
DLP	Endividamento de Longo Prazo	Contínua
END	Endividamento Total	Contínua
EST	Estrutura dos Ativos	Contínua
OC	Oportunidades de Crescimento	Contínua
DIMA	Dimensão da empresa em termos de ativo	Contínua
DIMVN	Dimensão da empresa em termos de volume de negócios	Contínua
ROA	Rendibilidade do ativo	Contínua
ROE	Rendibilidade do capital próprio	Contínua
ALV	Alavancagem	Contínua
LIQ	Liquidação/Risco	Contínua
Z	Z Score - Medida de Avaliação do Risco	Contínua
RENT	Rentabilidade da atividade operacional	Contínua
TEI	Taxa Efetiva de Imposto	Contínua
CFLOW	Cash-Flow	Contínua
TRAB	Número de Colaboradores	Contínua
PAÍS	Sede da Empresa	Nominal
TIPO	Tipo de Empresa: Micro, Pequena ou Média	Nominal
SETOR	Setor de Atividade	Nominal

Anexo 3 – Setores de Atividade

A – Agricultura, produção animal, caça, floresta e pesca

B - Indústrias Extrativas

C - Indústrias Transformadoras

D – Eletricidade, gás, vapor, água quente e fria e ar frio

E – Captação, trat. e distrib. de água; saneamento gestão de resíduos e despoluição

F - Construção

G – Comércio por grosso e a retalho; reparação de veículos automóveis e motocicletos

H – Transportes e armazenagem

I – Alojamento, restauração e similares

J - Atividades de informação e de comunicação

L - Atividades Imobiliárias

M – Atividades de consultoria, científicas, técnicas e similares

N – Atividades administrativas e dos serviços de apoio

O - Administração Pública e Defesa; Segurança Social Obrigatória

P - Educação

Q – Atividades de saúde humana e apoio social

R – Atividades artísticas, de espetáculos, desportivas e recreativas

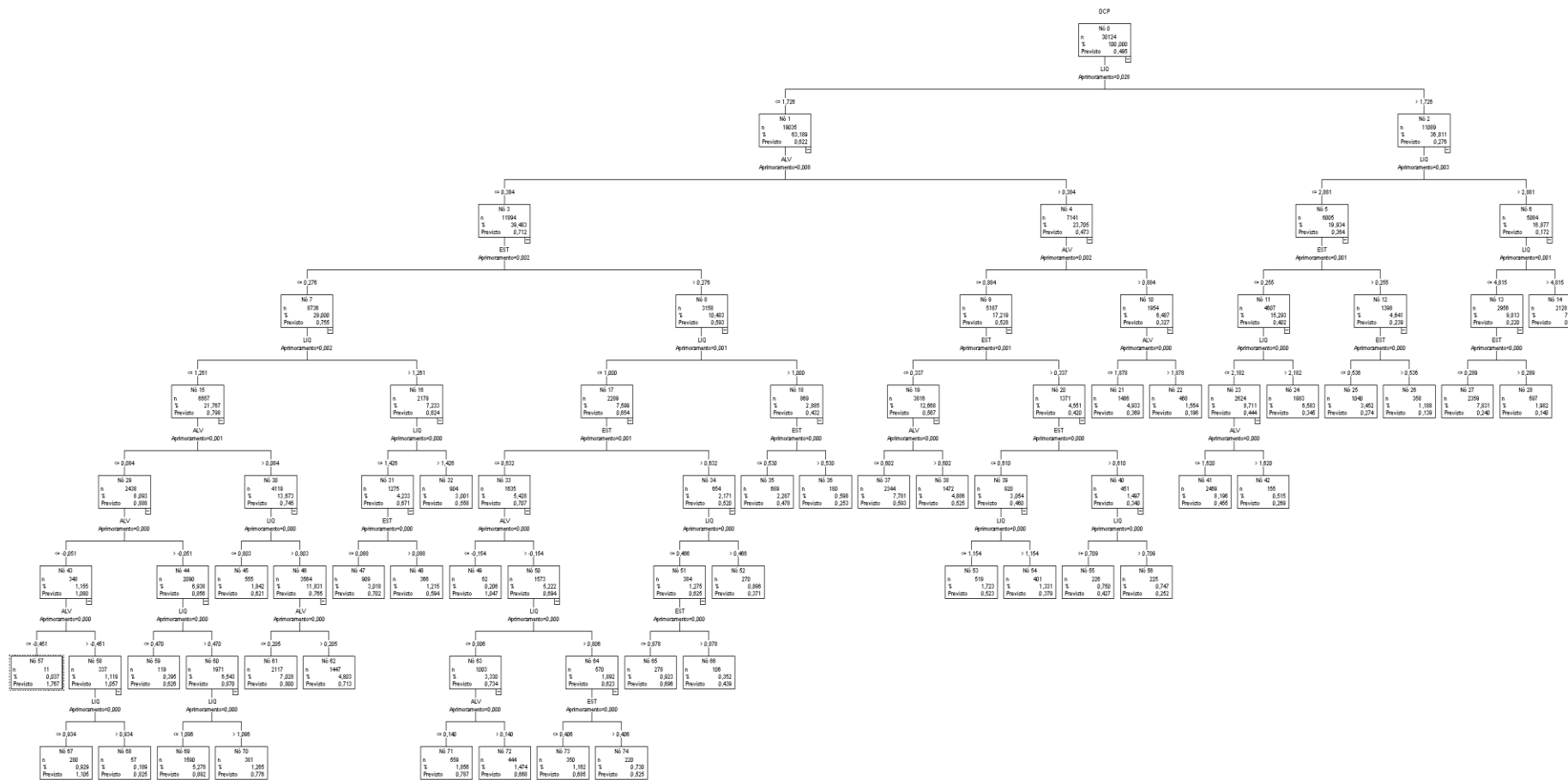
S – Outras Atividades de serviços

U – Atividades dos organismos internacionais e outras instituições extraterritoriais

Anexo 4 – Parâmetros da Árvore de Decisão - Maturidade de Curto Prazo

DCP		AD
Partição		50/50
Pré-Poda	Nº de Níveis	10
	Nº Mínimo de Observações por Nó pai	1%
	Nº Mínimo de Observações por Nó filho	0,5%
	Função diversidade	0,0001
Poda	Valor máximo da diferença do risco	Sim
Estrutura	Nº Nós folha	37
	Profundidade	8
Conjunto de Treino	R^2	83%
Conjunto de Treino	R^2	80%

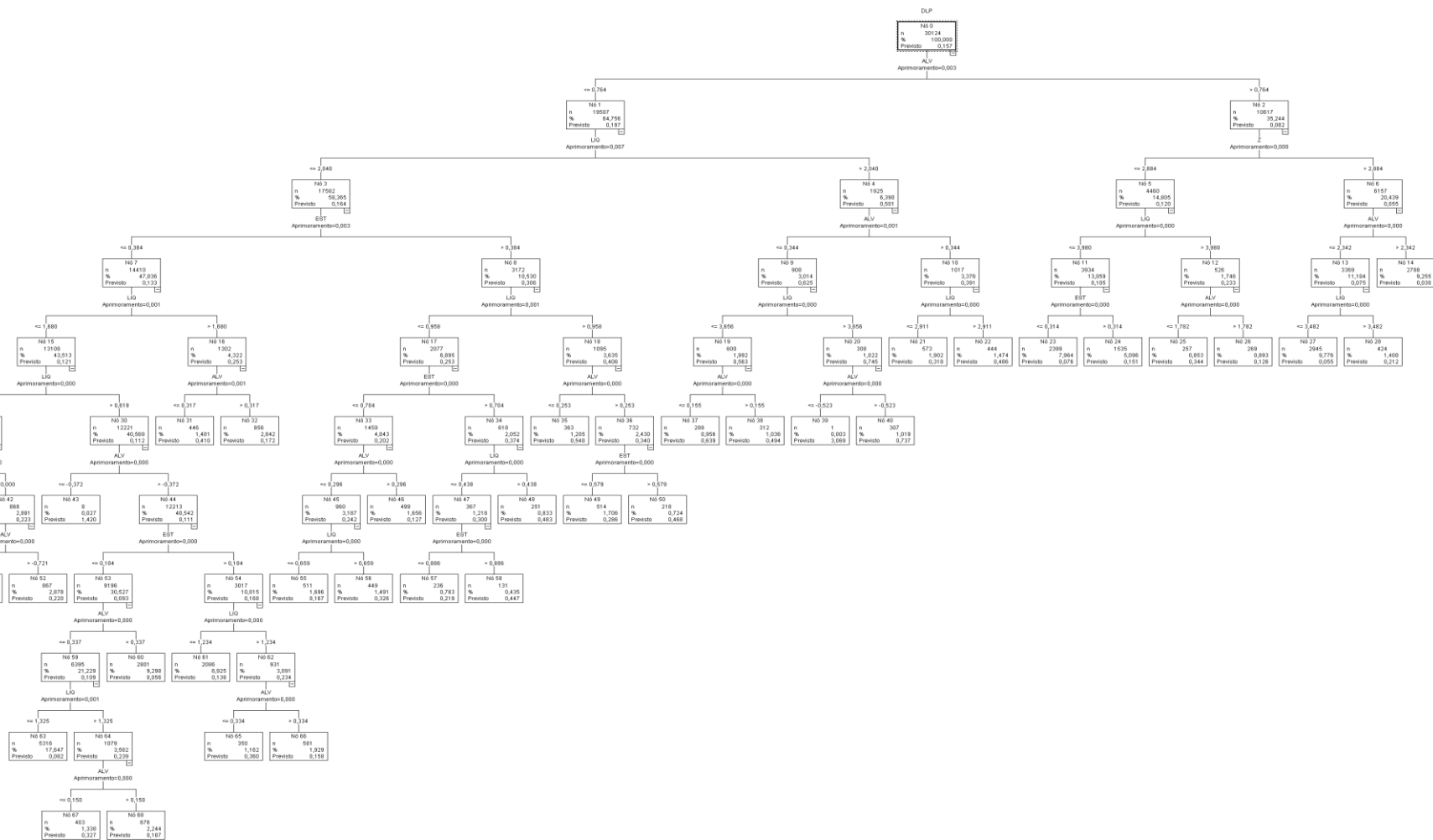
Anexo 5 – Árvore de Decisão – Maturidade de Curto Prazo



Anexo 6 – Parâmetros da Árvore de Decisão - Maturidade de Longo Prazo

DLP		AD
Partição		50/50
Pré-Poda	Nº de Níveis	10
	Nº Mínimo de Observações por Nó pai	1%
	Nº Mínimo de Observações por Nó filho	0,5%
	Função diversidade	0,0001
Poda	Valor máximo da diferença do risco	Sim
Estrutura	Nº Nós folha	32
	Profundidade	10
Conjunto de Treino	R^2	62%
Conjunto de Treino	R^2	61%

Anexo 7 – Árvore de Decisão – Maturidade de Longo Prazo



Anexo 8 – Parâmetros das Redes Neurais

	Partição	50/50
Opções de Computação	Nº Épocas	1000
	Otimização	Gradiente Descendente
	Função Ativação Camada Output	Tangente Hiperbólica
	Função Ativação Camada Intermédia	Tangente Hiperbólica
Parâmetros	Máximo Unidades Camada 1	Automático
	Máximo Unidades Camada 2	Automático

