



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO EM
CONTABILIDADE, FISCALIDADE E FINANÇAS
EMPRESARIAIS

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

ESTUDO DO INCUMPRIMENTO DA SITUAÇÃO
CONTRIBUTIVA E FISCAL DAS EMPRESAS UTILIZANDO
REDES NEURONAIAS

BRUNO MIGUEL DE MAGALHÃES SOARES MENDES

FEVEREIRO - 2017



LISBON
SCHOOL OF
ECONOMICS &
MANAGEMENT
UNIVERSIDADE DE LISBOA

MESTRADO EM
CONTABILIDADE, FISCALIDADE E FINANÇAS
EMPRESARIAIS

TRABALHO FINAL DE MESTRADO
DISSERTAÇÃO

ESTUDO DO INCUMPRIMENTO DA SITUAÇÃO
CONTRIBUTIVA E FISCAL DAS EMPRESAS UTILIZANDO
REDES NEURONAIAS

BRUNO MIGUEL DE MAGALHÃES SOARES MENDES

ORIENTAÇÃO:

PROF. DOUTOR EDUARDO BARBOSA DO COUTO

FEVEREIRO - 2017

Resumo

O objectivo principal desta dissertação de Mestrado prende-se com o desenvolvimento de um modelo de risco de incumprimento contributivo e fiscal das empresas, junto do Estado. Para tal utilizaram-se dados contabilísticos, financeiros e económicos das mesmas. Esta temática, de assaz importância para as empresas, investidores, trabalhadores, Estado e a consolidação da economia, não tem sido alvo de grande atenção nem por parte dos académicos, nem por parte da sociedade civil.

Na construção do modelo recorreu-se a uma técnica no domínio das Redes Neuronais, conhecida por *Multilayer Perceptron Backpropagation*, com o algoritmo de aprendizagem do gradiente descendente, através de uma amostra de 1000 empresas, sendo as variáveis independentes retiradas da base de dados da Informa D&B – nomeadamente balanço e demonstração de resultados, para o período de 2011 a 2015.

Tendo em conta os dados consolidados do grupo de treino e do grupo de teste, a percentagem de erro global do modelo, é de 6,6%, o que significa que o modelo não só é robusto e de qualidade, como também se comprova que o uso das Redes Neuronais tem um grande potencial, designadamente na generalização de padrões e em termos de capacidade preditiva.

Palavras-chave: Incumprimento contributivo e fiscal, risco, redes neuronais, *Multilayer Perceptron, Backpropagation*

Abstract

The main aim of this Master's thesis is the development of a Model of companies' tax and social security non-compliance risk. For this purpose, accounting, financial and economic data were used. This thematic of great importance for companies, investors, their workers, State and for the consolidation of the economy, has not generated until now a huge attention among academics and the civil society.

In the construction of the model we used a technique in the field of Neural Networks, known as Multilayer Perceptron Backpropagation with the gradient descending learning algorithm, moreover we delimited a sample of 1,000 firms in which the independent variables were taken from Informa D&B database (balance sheet and income statement) presented in the period from 2011 to 2015.

Taking into account the consolidated data of the training group and the test group, the overall error percentage of the model is 6,6%, which means that the model is not only robust but has quality, this also proves the use of Neural Networks has great potential, particularly in generalizing patterns and predictive capacity.

Keywords: Tax and social security non-compliance, risk, neural networks, Multilayer Perceptron, Backpropagation

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neuronais

Agradecimentos

Agradeço ao meu orientador, Prof. Eduardo Couto, o trilhar deste caminho.

Agradeço à Informa D&B pela disponibilização dos dados.

Agradeço aos meus amigos “imbecés” do risco.

Agradeço aos meus Pais e à minha prima Nela.

Agradeço à Lina, pelo que fostes, és, e serás... *“Without you I’m nothing”* – *Brian Molko*.

Lista de abreviaturas e siglas

AT – Autoridade Tributária e Aduaneira

BP – *Backpropagation*

CPPT – Código de Procedimento e de Processo Tributário

D.L. – Decreto - Lei

DM – *Data Mining*

DN – Despacho Normativo

GD – Gradiente Descendente

IVA – Imposto sobre o Valor Acrescentado

MLP – *MultiLayer Perceptron* ou Perceptrão Multicamadas

RN – Rede Neuronal ou Rede Neuronal Artificial

SP – Sistema Previdencial

SS – Segurança Social

Índice de tabelas

Tabela I – Caraterização da amostra final	23
Tabela II – Amostra de treino e amostra de teste	23
Tabela III – Informação RN MLP	24
Tabela IV – Resumo do modelo	25
Tabela V – Classificação da RN MLP	26
Tabela VI – Variáveis independentes iniciais*	35
Tabela VII – Importância das variáveis independentes	38
Tabela VIII - Estimativas de Parâmetros.....	40

Índice de imagens

Imagem 1 - Ligação de neurónios humanos.....	7
Imagem 2 - RN MLP - Camada de entrada, camada escondida e camada de saída 1	15
Imagem 3- RN MLP - Camada de entrada, camada escondida e camada de saída 2.	15
Imagem 4– Função tangente hiperbólica.....	16
Imagem 5– CRISP - DM.....	19
Imagem 6– Importância normalizada das variáveis independentes.....	25
Imagem 7 – Desenho da RN MLP.....	39

Índice

Capítulo 1-	Introdução	1
Capítulo 2-	Revisão da Literatura	3
	2.1- <i>Situação contributiva e fiscal</i>	3
	2.2 – <i>O passado e o presente das RN</i>	7
	2.3 - <i>Teses e trabalhos de RN</i>	10
	2.4 - <i>Trabalhos de fiscalidade e segurança social</i>	12
Capítulo 3-	Redes Neuronais – Teoria e Matemática	14
Capítulo 4-	Metodologia	19
	4.1 – <i>CRISP – DM</i>	19
	4.2 – <i>População, amostra e variáveis</i>	21
Capítulo 5-	Análise e Discussão de Resultados	24
Capítulo 6-	Conclusões	27
	Bibliografia	29
	ANEXOS	34
	ANEXO I – <i>Variáveis independentes iniciais</i>	35
	ANEXO II – <i>Importância das variáveis independentes</i>	38
	ANEXO III – <i>Desenho e cálculos da RN MLP</i>	39

Capítulo 1- Introdução

Uma empresa não é somente um ente de direito, representa uma unidade viva, que respira e se desenvolve pelas pessoas que a compõem e pelo meio em que se insere, bem como, pelo local onde atua. Para Chandler (1992), uma firma é uma entidade legal e administrativa na sua essência; depois de constituída, torna-se num conjunto de competências adquiridas, instalações e capital líquido. Finalmente, quando busca o lucro, torna-se no motor de desenvolvimento presente e futuro das economias modernas, através da produção e distribuição de bens e serviços.

Em termos contextuais, recentemente, a crise instalada a nível local, nacional e internacional (designada por crise do *subprime* e da dívida soberana¹), e a posterior chegada do FMI em 2011, conduziu o nosso país a um período marcado por medidas de forte austeridade, a par de uma tentativa de recuperação da economia portuguesa. Todo o tecido empresarial lidou com a crise, mais concretamente desde 2008/2009 de uma maneira diferenciada. Muitas empresas faliram, outras fecharam, algumas tentaram a internacionalização, outras despediram pessoas para conter custos, *i.e.*, cada qual tentou à sua maneira lidar com a crise. E qual a situação das empresas face aos pagamentos ao Estado, à Segurança Social (SS), à Autoridade Tributária e Aduaneira (AT)? Seria lógico que as empresas não insolventes suprissem primeiro os pagamentos dos impostos, taxas e contribuições legalmente obrigatórias, uma vez que o carácter coercivo do incumprimento pode levar a penhoras, arrestos e dificuldades no seu normal funcionamento.

Neste contexto, emerge desde logo a seguinte questão de investigação: olhando ao comportamento das empresas determinado período temporal, será possível saber que empresas têm uma tendência para não cumprir as suas obrigações com o País?

¹ Originária nos Estados Unidos da América, decorrente da concessão de empréstimos hipotecários de alto risco, as ondas de choque levaram a uma recessão económica internacional e aceleraram os problemas na zona euro, nomeadamente na Irlanda, Grécia e Portugal – para maior profundidade ver Pinto (2014).

Neste âmbito, pretende-se classificar, detetar e analisar as empresas que possam incumprir com as suas obrigações contributivas e fiscais, quando esse montante teria que ser liquidado nos cofres do Estado.

Este estudo é inovador, uma vez que existem poucas publicações, dissertações ou teses de mestrado e/ou doutoramento sobre o comportamento das empresas relativamente ao incumprimento perante o Estado. Espera-se por um lado, aprender com os seus resultados e por outro, contribuir para que se consiga detetar corretamente e de forma prospectiva quais os futuros devedores. Cada contribuinte incumpridor significa menos eficácia fiscal e maior fuga aos impostos; ter dívidas representa um sinal de alarme do comportamento presente e futuro da entidade empresa e do seu envolvimento na sociedade e na recuperação da economia.

Ao nível da metodologia foi utilizada a CRISP-DM para o tratamento dos dados, e foi usado a RN *MultiLayer Perceptron* (MLP) com o algoritmo *Backpropagation* (BP) para construção do modelo, alicerçados no tratamento informático do IBM SPSS *Modeler* e *Statistics*. As bases de dados utilizadas foram graciosamente cedidas pela Informa D&B, e abarcam um conjunto de dados (contabilísticos e outros), desde o ano de 2011 a 2015.

A percentagem de identificação global do modelo é de 93,4%, comprovando-se que a partir dos dados existentes é possível obter um modelo com qualidade, alcançando o objetivo proposto.

Em termos organizativos, esta dissertação estrutura-se em 6 capítulos, sendo que o capítulo seguinte apresenta a revisão da literatura científica mais relevante sobre o tema em análise. O capítulo 3 explica o conceito de Redes Neuronais (RN) e o seu tratamento matemático. A metodologia de investigação e delimitação da amostra está patente no capítulo 4. A apresentação, discussão e interpretação dos resultados aparecem no capítulo 5. No último capítulo são formuladas as conclusões e apontadas pistas para investigações futuras e possíveis implementações deste modelo.

Capítulo 2- Revisão da Literatura

2.1- Situação contributiva e fiscal

Uma empresa tem como objetivo final o lucro / criação de valor. No entanto, as particularidades de cada sector de actividade e cada empresa, tornam fascinante a análise do caminho que tomam para a persecução do objetivo.

Mais que analisar rácios económicos e financeiros, convém ver algumas particularidades evidentes nas demonstrações financeiras e contabilísticas, tal como está definida na Norma Contabilística de Relato Financeira (NCRF 1), utilizando o balanço (estado financeiro da empresa) e a demonstração de resultados (resumo financeiro dos resultados das operações da empresa). O foco deste trabalho é no incumprimento da situação contributiva e fiscal, logo abrange as dívidas ao Estado, designadamente à Segurança Social (SS) e à Autoridade Tributária e Aduaneira (AT).

Neste estudo, não interessa pormenorizar o conceito de empresa, sendo aqui importante realçar que na base da sua existência tem um número de identificação fiscal coletivo e um CAE (classificação de atividades económicas)².

Para Páscoa (2009, p.2), a SS “enquadra num sistema de políticas económicas e sociais, cujo objeto de atuação ultrapassa o indivíduo, englobando a sociedade em que este se encontra inserido, tendo em conta a situação social

² Tal como descrito em <http://www.sicae.pt/>, O SICAE contém informação sobre o código CAE das seguintes pessoas coletivas e entidades equiparadas:

- a. Sociedades comerciais e civis sob forma comercial, cooperativas, empresas públicas, agrupamentos complementares de empresas e agrupamentos europeus de interesse económico;
- b. Estabelecimentos individuais de responsabilidade limitada;
- c. Representações de pessoas coletivas com sede no estrangeiro que habitualmente exerçam atividade em Portugal e estejam sujeitas a registo comercial;
- d. Instrumentos de gestão fiduciária e sucursais financeiras exteriores registados na Zona Franca da Madeira;
- e. Quaisquer outras pessoas coletivas sujeitas a registo comercial

de cada indivíduo, os riscos sociais e as medidas necessárias a adoptar para cobrir esses riscos”. Na consulta do seu site institucional, em <http://www.seg-social.pt/objectivos-e-principios> , é sublinhado alguns dos seus princípios orientadores, tais como, assegurar dos direitos básicos, assim como, a igualdade de oportunidades entre todos, tendo como finalidade o bem-estar e da coesão social para toda a população residente ou trabalhadora no nosso País.

Atualmente, a Lei de bases gerais do sistema de Segurança Social, Lei n.º 4/2007 de 16 de janeiro, regula o seu funcionamento. O sistema da SS está dividido em três (3) partes: Sistema de Protecção Social de Cidadania, Sistema Previdencial (SP), e, Sistema Complementar³.

As empresas têm um papel importante no financiamento da SS, e concretamente no SP. No artigo 50º, definem-se os objetivos do SP, nomeadamente assegurar “prestações pecuniárias substitutivas de rendimentos de trabalho perdido em consequência da verificação das eventualidades legalmente definidas”. O artigo 56º - Obrigações dos contribuintes e o artigo 59º - Responsabilidade pelo pagamento das contribuições, são da maior importância para a tipificação de quem será o devedor do incumprimento à SS.

Para concluir, importa fazer referência ao artigo 74º - Certificação da regularização das situações. O n.º 1 diz que pode ser requerido em qualquer altura, declaração comprovativa do regular cumprimento das obrigações.

A AT, conforme o inscrito no D.L. n.º 118/2011 de 15 de dezembro, artigo 2º, n.º1, tem por missão, “administrar os impostos, direitos aduaneiros e demais tributos que lhe sejam atribuídos, bem como exercer o controlo da fronteira externa da União Europeia e do território aduaneiro nacional, para fins fiscais, económicos e de protecção da sociedade, de acordo com as políticas definidas pelo Governo e o Direito da União Europeia”. De entre as suas atribuições, salienta-se:

³ Para melhor entendimento das teorias subjacentes, ver Pascoa (2009) e Azevedo (2014).

- Assegurar a liquidação e cobrança dos impostos, direitos aduaneiros e tributos que administra, assim como outras receitas do Estado ou de pessoas colectivas de direito público;
- Exercer a ação de inspecção tributária e aduaneira;
- Exercer a ação de justiça tributária;
- Promover a correta aplicação da legislação e das decisões administrativas;
- Informar os contribuintes das suas obrigações fiscais e aduaneiras, apoiando-os no seu cumprimento.

Para a AT, a definição da Situação Tributária Regularizada, está inscrita no Código de Procedimento e de Processo Tributário (CPPT), diretamente no artigo 177^o-A, ao dizer que tal implica não dever impostos ou outras prestações tributárias, assim como os respectivos juros de mora; autorização de pagamento em prestações, com garantia prestada nos termos legais; outras situações com garantia constituída nos termos legais, designadamente a intenção, ou a apresentação de meio gracioso, contencioso ou judicial de discussão das dívidas.

Curiosamente, do Programa SIMPLEX 2006, surgiu o D.L. n.º 114/2007 de 19 de abril, que dispensa a apresentação de certidões comprovativas da situação contributiva e fiscal regularizada, por parte de quem participe em procedimentos administrativos tipificados no referido diploma. Para reforçar esta mudança, foi introduzido no CPPT o artigo 177^o-C – Comprovação de situação tributária.

O que a SS considera como situação contributiva regularizada, está plasmado no Código dos Regimes Contributivos do Sistema Previdencial de Segurança Social, Lei n.º 110/2009 de 16 de setembro e posteriores alterações, mais concretamente o artigo 208^o, que no n.º 1 refere a inexistência de dívidas de contribuições, quotizações, juros de mora e de outros valores, e no n.º 2, outras situações abrangidas, como o pagamento em prestações autorizado e o cumprimento das condições implícitas. Igualmente, no caso de reclamação,

recurso, dedução de oposição ou impugnação judicial da dívida, desde que, constituindo garantia ou dispensa da mesma, no estrito cumprimento da lei.

O legislador entendeu, no caso da parte fiscal, reforçar o que a não regularização da situação tributária significa. Assim, o artigo 177º-B do CPPT refere ser vedados aos contribuintes:

- Celebrar ou renovar contratos de fornecimento, empreitadas de obras públicas ou aquisição de serviços e bens com o Estado, regiões autónomas, institutos públicos, autarquias locais e instituições particulares de solidariedade social;
- Concorrer à concessão de serviços públicos;
- Cotar em bolsa de valores os títulos do seu capital social;
- Lançar ofertas públicas de venda do seu capital ou vender em subscrição pública, títulos de participação, obrigações ou acções;
- Ter apoios de fundos europeus estruturais e de investimento e públicos;
- Distribuir lucros ou fazer adiantamentos no decurso do exercício;

Por outro lado, no Estatuto dos Benefícios Fiscais, o artigo 13º refere que quem não pague qualquer imposto e contribuições devidas à SS, fica impedido do reconhecimento do direito a benefícios fiscais. Ainda no mesmo registo, o Código Fiscal do Investimento, D.L. n.º 162/2014 de 31 de outubro, apenas é para o promotor que tiver a situação contributiva e tributária regularizada.

Muitas empresas ficam com a noção de que é mais fácil não pagar as obrigações contributivas e fiscais do que as suas contas do dia-a-dia. Por outro lado, a existência de vários planos especiais de regularização de dívidas, desde o D.L. 124/96 de 10 de agosto – Plano Mateus, até ao recente D.L. n.º 67/2016 de 3 de novembro – Programa Especial de Redução do Endividamento ao Estado (PERES), sugere às empresas que o Estado, em qualquer altura que necessite, abre mão de alguma parte da dívida existente, e assim, alivia o total dos montantes a pagar.

2.2 – O passado e o presente das RN

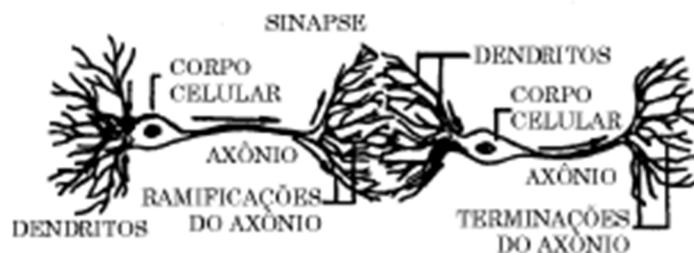
De seguida iremos efectuar uma breve resenha histórica sobre a evolução das RN, acompanhada de uma pequena explicação sobre os seus conceitos básicos.

O comportamento do homem e em especial o funcionamento do cérebro e as funções do sistema nervoso intrigou desde tempos remotos a humanidade e levou a inúmeros estudos sobre o assunto. O cérebro, a sua interação e conectividade com o restante corpo e a maneira como este o controla e define, revelam a sua complexidade.

A definição do que constitui o neurónio, tal como observada na imagem 1, com maior ou menor complexidade, é de aceitação geral, convém no entanto, lembrar que é impossível reproduzir com exactidão todos os processamentos humanos, uma vez que nem sequer conseguimos saber o número exacto de neurónios.

A figura infra representa basicamente dois neurónios humanos em ligação:

Imagem 1 - Ligação de neurónios humanos⁴



Fonte: <http://www.biomania.com.br/bio/?pg=artigo&cod=1243>, acesso em 24-09-2016

Da curiosidade, da vontade e dos estudos, e também do engenho do homem, aproveitando a evolução das máquinas, tentou-se reproduzir o funcionamento do cérebro humano através da criação de modelos computacionais, ou redes artificiais, que foram designadas por Redes Neuronais ou Redes Neuronais Artificiais (RN). Num registo mais simples, uma RN, é um modelo inspirado no comportamento do sistema nervoso, de modo a criar uma inteligência artificial

⁴ Neurónio - é a célula do sistema nervoso responsável pela condução do impulso nervoso.

Axónio – Filamento de transmissão de sinais.

Dendrite – Zona recetora.

Sinapse – Unidade estrutural e funcional onde os neurónios interagem entre si.

similar à humana. Para Haykin (2009, p. 2) tal, “ é um processador maciçamente paralelamente distribuído constituído por unidades de processamento simples, que têm uma natural propensão para armazenar conhecimento experimental, tornando-o disponível para uso. Assemelha-se ao cérebro em dois aspectos:

1. O conhecimento é adquirido pela rede através de um processo de aprendizagem.
2. Forças de conexão entre neurónios, conhecidas como pesos sinápticos, são utilizados para armazenar o conhecimento adquirido. ”

Segundo Bar (2014, p. 185), as “ Redes Neurais Artificiais são algoritmos e técnicas que podem ser usadas para efectuar modelos de estatística não-linear e fornecer uma nova alternativa para a regressão logística (...)”. Como se depreende, as definições de RN são inúmeras e complementares. Interessa agora entender a sua evolução e desenvolvimento através de uma breve descrição cronológica de acontecimentos relevantes nas RN.

McCulloch and Pitts (1943) foram os pioneiros na perceção das funções do sistema nervoso, criando um modelo matemático para uma RN.

Em 1949, Hebb, trabalhou numa teoria geral do comportamento, com regras de aprendizagem para RN, propondo uma lei de aprendizagem específica para as sinapses dos neurónios.

Até esta data, os modelos propostos diziam respeito à aprendizagem não supervisionada, ou seja situações em que o output desejado não é previamente conhecido e conseqüentemente não existe orientação no processo de treino.

Em 1958, Rosenblatt propôs o Perceptrão como o primeiro modelo de aprendizagem supervisionada. Tal é concebido como uma estrutura com uma camada de neurónios ligados aos inputs do sistema. As ligações têm pesos numéricos que podem ser mudados durante o treino, de acordo com as regras de aprendizagem, uma vez que (em vez de maneira a que) para cada *input* X na rede, o *output* Y verdadeiro também é conhecido (em vez de é dado).

Deste trabalho pioneiro surgiram outros modelos baseados no Perceptrão, como por exemplo, o LMS (Algoritmo *Least-Mean-Square*) e a rede Adaline

(*ADaptative LInear Network*) , de Widrow e Hoff (1960)(Haykin, 2009; Rojas, 1996).

No entanto, na década de 60, começaram as críticas ao Perceptrão, que tiveram em Minsky e Papert os seus maiores detratores. Em 1969, conseguiram provar com uma base matemática sólida as limitações computacionais do Perceptrão de um nível, designadamente a sua incapacidade de fazer generalizações globais partindo de exemplos localmente aprendidos. A aceitação deste trabalho e das suas conclusões, levou a uma diminuição de verbas monetárias disponibilizadas para investigação, desaceleração nas pesquisas e abandono das teorias subjacentes às RN (Minsky & Papert, 1969).

No decurso da década de 80, com o desenvolvimento de um maior poder computacional, foi registado um incremento relevante no estudo e aplicação das RN. São exemplo, os trabalhos de Carpenter e Grossberg e a *Adaptative Resonance Theory* (ART), desenvolvida desde 1976, que foi evoluindo como uma série de modelos de redes neurais em tempo real, que executa aprendizagem com e sem supervisão, reconhecimento de padrões e previsão (Carpenter & Grossberg, 2014).

Em 1982, surgiram as redes de Hopfield - recorrentes, com uma função de energia e onde cada neurónio está ligado a todos os outros, bem como, os Mapas auto-organizados de Kohonen – sem supervisão, ou seja, os neurónios organizam-se de acordo com determinados valores de *input* (Hopfield, 1982; Kohonen, 1982) .

Introduzido por Paul Werbos em 1974, o algoritmo de *Backpropagation* (BP) conduziu a um movimento liderado por Rumelhart (1986), e que ainda hoje se mantém ativo e em força. Este é um método de aprendizagem supervisionado e que provou ser computacionalmente eficiente no treino de uma RN (Perceptrão) de várias camadas, designado por *Multilayer Perceptron* (MLP) (Baidyk et al, 2010; Chauvin & Rumelhart, 1995).

Em 1992, Boser, Guyon e Vapnik desenvolvem as *Support-vector machines*, que representam um avanço em relação aos algoritmos tradicionais usados na

seleção de pesos e limites. O seu resultado é um classificador que tende a ser mais exato em dados invisíveis (Leskovec et al., 2014).

No início deste milénio, e até aos nossos dias, as RN tornaram-se ainda mais estudadas, devido ao despontar do *Deep Learning* - um ramo de *Machine Learning*, baseado num conjunto de algoritmos para aprender múltiplos níveis de representação a fim de modelar relações complexas entre os dados (Deng & Yu, 2014).

Em Portugal, ao nível académico, tem-se acompanhado a evolução das RN até ao *Deep Learning*, e como exemplo disto, no Instituto Superior Técnico (IST) organiza-se desde 2011 a *Lisbon Machine Learning School* com especialistas portugueses e internacionais para durante 8 dias trocarem experiências.

2.3 - Teses e trabalhos de RN

O uso de RN em estudos, teses e múltiplas publicações, massificou-se nas três últimas décadas, não tanto por estar na moda, mas mais porque as mesmas estão voltadas para a resolução de problemas de uma forma elegante e prática, chegando a resultados conclusivos.

A ideia subjacente é de que, ainda que a pesquisa e o desenvolvimento de RN continue, as necessidades da sociedade e os desafios do planeta conduzem a um carácter prático e não teórico. A criação de *softwares* como o IBM *Neural Network Utility* (1994) e o SAS *Enterprise Miner* (1997), com as constantes atualizações e melhorias introduzidas até aos nossos dias, faz com que a utilização de RN esteja ao dispor de universidades, grandes empresas e governos, liderando assim a nova era do *Big Data* e do *Data Mining* (DM) (Gupta & Smith, 2000).

Wong, Bogdanovich e Selvi (1997), analisaram as aplicações de RN no setor dos negócios empresariais no período entre 1985 e 1995, tendo sido aceites 203 artigos. A metodologia usada centrou-se na Base de Dados ABI/INFORM, que providenciou acesso a mais de 800 publicações mundiais. Igualmente, foi efectuado uma pesquisa manual da *Business Periodical Index* (BPI), de livros de ensino sobre RN e 15 revistas. Em termos percentuais, as áreas mais

relevantes são: produção/operação, com 53,5 %; finança, 25,4 % e sistemas de Informação, com 7,5 %.

Wong e Selvi (1998) analisaram as aplicações de RN no sector da Finança no período entre 1990 e 1996, tendo considerado como aceites 64 artigos. A metodologia usada centrou-se na Base de Dados ABI/INFORM, que providenciou acesso a mais de 800 publicações mundiais. Igualmente, foi efetuada pesquisa manual da *Business Periodical Index* (BPI), de livros de ensino sobre RN e 12 revistas. As áreas com mais artigos publicados são: previsão de falências de empresas – 8; previsão de desempenho/seleção de ações – 6; previsão de falências de bancos / instituições de poupança – 5.

Do artigo de Wong, Lai e Lam (2000), acerca do período 1994-1998, e na aplicação de RN ao mundo empresarial, é de referir que dentro dos 302 artigos selecionados, em 1995 (76 artigos), 1996 (78) e 1997 (69), representam 73,8 % do total, e que em relação ao estudo de 1997, é curioso constatar o aumento dos estudos em marketing/distribuição e sistemas de informação, enquanto a Produção/Operação e Finanças continuam a ser as mais focadas. Os autores perspetivam que no futuro as RN vão ter um interesse redobrado por parte dos académicos e profissionais, uma vez o desenvolvimento informático e a possibilidade de aplicar as RN à resolução de problemas práticos, abre uma nova via de investigação.

Paliwal e Kumar (2009) apresentaram indicações úteis sobre a capacidade das redes neurais e dos modelos estatísticos usados nas diferentes aplicações, através da sua comparação. Dividindo em categorias, designadamente, contabilidade e finança, medicina e saúde, engenharia e manufatura, marketing, e, aplicações gerais, e analisando qual a melhor performance. Globalmente, as RN obtiveram uma performance melhor ou tão boa como os outros métodos, no entanto, a maioria dos artigos não utilizaram técnicas estatísticas optimizadas.

Tkac e Verner (2016), analisaram a aplicação das RN ao mundo empresarial entre 1994 e 2015. Dos 412 artigos selecionados, dificuldades financeiras e falências com 75, e, ações e títulos com 73 artigos, foram as áreas mais estudadas. O tipo de RN mais utilizado foi a *Multilayer Perceptron feedforward*

(MLP) e o algoritmo de aprendizagem mais usado foi o gradiente descendente (GD).

2.4 - Trabalhos de fiscalidade e segurança social

Bench-Capon (1993), utilizou um problema ficcional de pagamento de subsídios de SS a idosos, como forma de demonstrar que as RN podem classificar os problemas com sucesso, podem ser catalisadoras para a correta identificação e criação de regras que descrevam os problemas propostos. As fases de treino e teste foram criados com o programa LISP (treino com 2400 observações, 64 inputs), foram construídas RN com uma, duas e três camadas escondidas, usando o programa “*The Aspirin*”, que obtiveram respetivamente taxas de sucesso de, 99,25%, 98,90%,e, 98,75%. O autor concluiu, genericamente, que as RN têm uma performance e classificação boas, embora dependa da qualidade da fase de treino, da percentagem de utilização de casos limite e da correta identificação dos inputs.

Hong e Hong (2007), apresentaram um estudo sobre um sistema preventivo do risco de investimento no fundo da SS na R.P.China. Utilizaram nesta análise uma RM MLP BP, com 10 inputs iniciais, uma camada escondida, e 3 indicadores na camada de saída, para preverem o grau de aviso indicativo do risco. Da análise aos resultados, concluiu-se que em face do elevado risco de rutura do sistema de SS, estudos e testes com RN, podem ser muito vantajosos.

Dentro dos estudos sobre tributação e assuntos fiscais, Denton e al. (1995), debruçaram-se sobre a categorização dos trabalhadores para efeitos fiscais. Utilizaram RN e modelos de regressão, bem como o método tradicional, partindo de um total de 120 casos, e dentro de 11 atributos escolheram 5. A conclusão é que os modelos utilizados ultrapassam os métodos tradicionais de avaliação dos casos.

Cushing Jr. e Arguea (1997) abordaram o mesmo tema, chegando a um resultado de acerto global de 97,2% nos dados de amostra, e ao efetuarem a classificação pela regressão logística atingiram resultados semelhantes.

Shaoqiu e Yingying (2010) utilizaram a RN *BP* para estabelecer um modelo de previsão de impostos que minimize o problema do *Overfitting*, utilizando os boletins estatísticos da China, do ano de 1981 a 2001, e 3 variáveis independentes (PIB, gastos totais e depósitos de poupança).

Chen et al. (2011) propuseram a aplicação de modelos de RN para detetar relatórios fiscais errados no setor da construção civil, especificamente para a região norte de Taiwan, usando 5769 relatórios de 3172 firmas. Concluiu-se que em qualquer modelo de RN utilizado, a taxa de acerto rondava os 80%.

González e Velásquez (2013) publicaram um artigo para caracterizar e detetar os contribuintes que usam faturas falsas num dado ano, usando diferentes tipos de DM, incluindo as RN. Este estudo tem por base o trabalho feito por Autoridades Tributárias Europeias e Americanas e apoiará a Administração Fiscal Chilena. As RN são usadas nos Estados Unidos, no Canadá, no Reino Unido, na Bulgária, no Peru e Chile. Dentro dos modelos de RN usados, o MLP foi o modelo com melhor performance, com 92% de acerto na deteção de fraudes nas pequenas empresas, e, de 89% nas médias e grandes empresas.

Stanislav Cut (2015) utilizou RN, árvores de decisão e *random forests*, para avaliar a capacidade de classificação destes métodos de DM, dentro da avaliação do risco das entidades sujeitas a Imposto sobre o Valor Acrescentado (IVA), na Eslováquia. Foram utilizados os dados fiscais do ano de 2012 dos contribuintes obrigados a entregar declarações de IVA, recorrendo a 26 variáveis explicativas focadas na avaliação do risco. O modelo de RN utilizado foi o MLP, com BP e utilizando o algoritmo de aprendizagem do GD. De salientar que o MLP foi o modelo que apresentou melhor capacidade de classificação.

Capítulo 3- Redes Neurais – Teoria e Matemática⁵

Como referido anteriormente, as RN são sistemas de processamento de informação, compostas por vários elementos computacionais simples que interagem entre si. Inspirados na arquitectura do cérebro humano, estes sistemas têm a capacidade de captar padrões complexos e de os aplicar a novos conjuntos de dados.

Cada elemento computacional da rede neural é chamado de neurónio e comunica com os demais por meio de conexões diretas.

De forma simplista, pode dizer-se que um neurónio recebe informação de outros neurónios (inputs ou sinais). Com base em aprendizagens anteriores o neurónio atribui diferente importância à informação recebida, aplicando pesos aos sinais (pesos sinápticos). Esta informação é processada e finalmente é tomada uma decisão transmitida aos neurónios seguintes.

Assim, o funcionamento dos neurónios pode ser descrito em duas fases: na primeira é calculada a soma dos sinais de entrada, atribuindo pesos aos sinais (pesos sinápticos); na segunda, aplica-se, à soma dos sinais, uma função de saída, chamada de função de ativação.

Uma RN pode encarar-se como uma sucessão/sequência de camadas de neurónios. A primeira camada é a camada de entrada e abarca as variáveis de entrada do modelo. O resultado de uma camada torna-se o início da camada seguinte, até à camada de saída. A função de ativação é a transformação perpetrada em cada fronteira entre camadas.

Concretamente, neste trabalho utilizamos uma RN MLP. O MLP não é mais que uma RN com uma ou mais camadas ocultas. É consensualmente definido como o tipo de RN mais usado pelos investigadores, essencialmente para criar modelos de classificação, tal como é o pretendido neste estudo, indo assim ao encontro do nosso objetivo.

Segundo Hayken (2009), o MLP assenta em 3 características básicas: a RN tem uma função de ativação não linear que é diferenciável; a RN tem uma ou mais

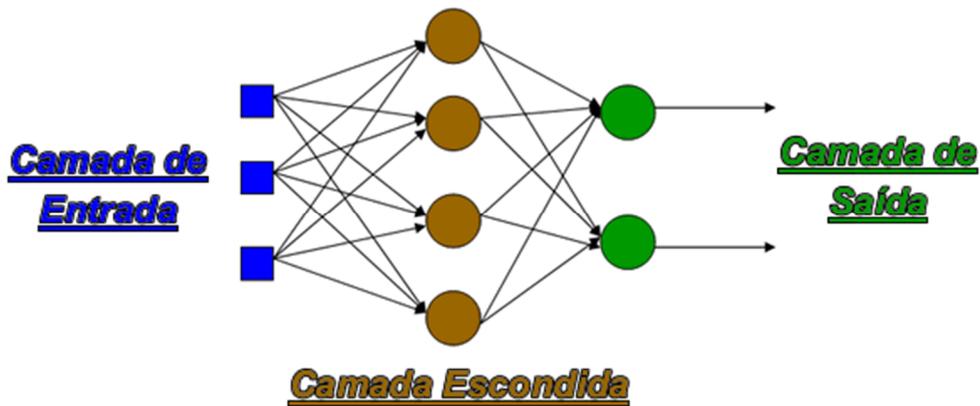
⁵ A teoria matemática subjacente às RN baseou-se em Haykin (2009), Bishop (2006) e Rojas (1996).

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neurais

camadas escondidas entre as camadas de entrada e saída; a RN possui um elevado grau de conectividade (determinada pelos pesos sinápticos).

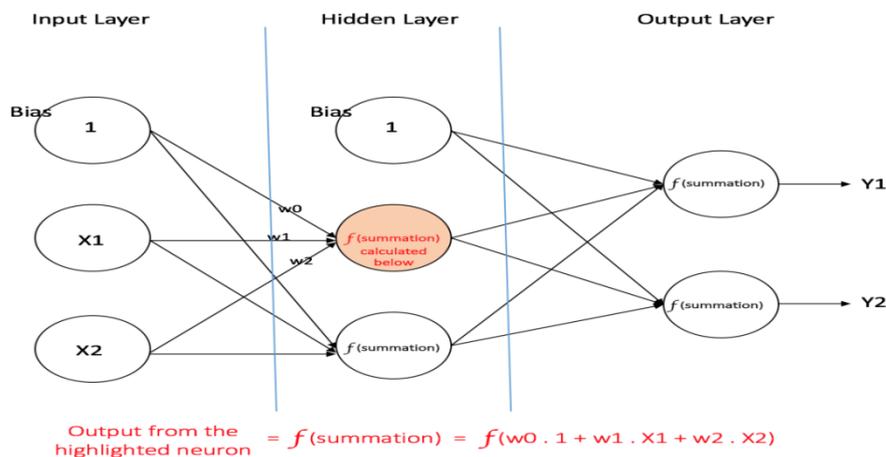
Imagem 2 - RN MLP - Camada de entrada, camada escondida e camada de saída 1



Fonte: <http://web.tecnico.ulisboa.pt/ana.freitas/bioinformatics.ath.cx/bioinformatics.ath.cx/index9>

bb2.html?id=111, acesso em 09-07-2016

Imagem 3- RN MLP - Camada de entrada, camada escondida e camada de saída 2



Fonte: <https://ujwlkarn.files.wordpress.com/2016/08/ds.png?w=1128>, acesso em 31-07-2016

Matematicamente, pode descrever-se o processo de funcionamento num neurónio n da camada k , da seguinte forma: é preciso somar o contributo de todos os neurónios da camada $k-1$; cada neurónio da camada $k-1$ contribui com um valor numérico e um peso. Um neurónio é um número real. Formalmente, se tivermos m neurónios contributivos da camada $k-1$,

$$n = \sum_{j=1}^m w_{kj} x_j$$

onde w_{kj} é um peso sináptico e x_j é o valor contributivo de um neurónio da camada $k - 1$, ou sinal de entrada. Se representarmos por φ esta transformação inter-camadas temos que:

$$y = \varphi(n)$$

A função φ toma o nome de função de ativação.

Esta função de ativação serve para determinar o valor de saída do neurónio, limitando o intervalo permitido de amplitude do sinal de saída a um valor finito, sendo desejável que seja diferenciável. As funções não-lineares mais usadas são as funções sigmoidais (funções em forma de S alongado).

A função tangente hiperbólica é a mais eficiente das funções sigmodais, em virtude da sua simetria, com uma variação entre -1 e 1 – o que facilita a determinação de máximos e mínimos pelo método do gradiente.

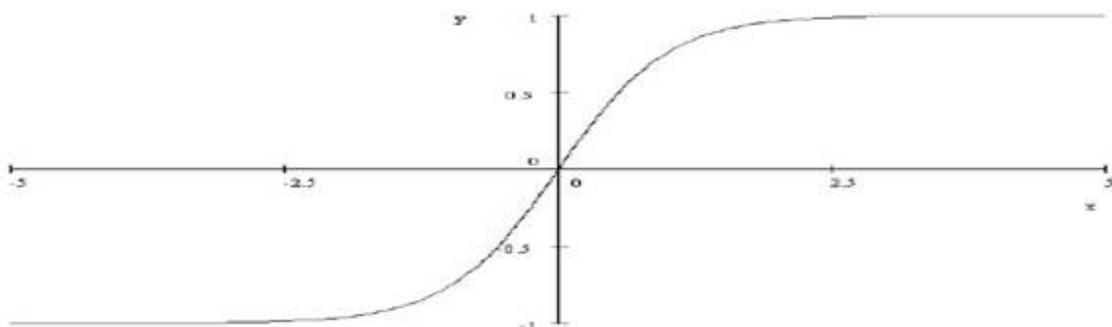
Convém aplicar uma transformação afim à saída de n :

$$v = n + b$$

Ficamos então com $y = \varphi(v)$.

A utilização do viés (*bias*) b , é essencial para o funcionamento da RN. Normalmente assume o valor de 1, e serve para ultrapassar problemas decorrentes dos valores de inicialização das camadas serem zero. O viés está ligado a cada neurónio de cada camada.

Imagem 4– Função tangente hiperbólica



Fonte: www.roperld.com, acesso em 11-06-2016

$$\varphi(v) = \tanh(v) = \frac{e^v - e^{-v}}{e^v + e^{-v}}$$

Assim, ao atingirmos a camada de saída desejada, estamos a utilizar a função de ativação, tantas vezes quantas as camadas anteriores. Se tivermos j camadas, temos que contemplar a atuação de $j-1$ vezes da função de activação, i.e.

$$y_{k \text{ saída}} = \varphi^{j-1}(v_k)$$

Note-se que $\varphi^0 = id$, função identidade.

A arquitetura apresentada no MLP é a *feedforward*, que como o próprio nome indica, “alimentação para a frente”, i.e., entre a camada de entrada (*input*) e a camada de saída (*output*) não existe qualquer interligação (nem ciclos nem voltas).

Basicamente, o algoritmo de BP funciona em 2 fases, a *forward* e a *backward*. Na primeira fase, e de acordo com a RN *feedforward*, após a fixação dos pesos sinápticos existe a propagação do sinal da camada de entrada pela RN, camada a camada, até atingir a camada de saída. Na fase *backward*, a comparação da camada de saída com a resposta pretendida, gera um sinal de erro. Este sinal de erro, é propagado pela RN, camada a camada, mas na direção da camada de entrada, promovendo ajustes sucessivos nos pesos sinápticos da RN.

O sinal de erro não é mais do que

$$e_j(k) = d_j(k) - y_j(k)$$

E representa a diferença entre o valor da real resposta pretendida e o valor gerado na camada de saída.

Recorremos ao algoritmo BP como método de aprendizagem. O BP é o algoritmo mais usado desde a sua criação na década de 80, sendo um algoritmo de aprendizagem supervisionado, onde são fornecidos os valores de entrada e os pretendidos na saída (p.e. Rojas, 1996), e usa o método do GD para minimização do total dos erros quadráticos.

Nesta presente dissertação, a escolha das RN assentou em 2 ideias basilares.

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neuronais

Por um lado, as RN representam uma abordagem que realiza boas previsões para dados gerados em processos que apresentam mudanças repentinas e que são conseqüentemente não lineares. Dada a volatilidade inerente aos dados económicos e financeiros (fortemente influenciados por mudanças governamentais, alterações legislativas e conjunturas internacionais), esta característica é fundamental. Por outro, a relativa simplicidade e flexibilidade na resposta a problemas difíceis e intrincados e dependências complexas ao nível dos dados.

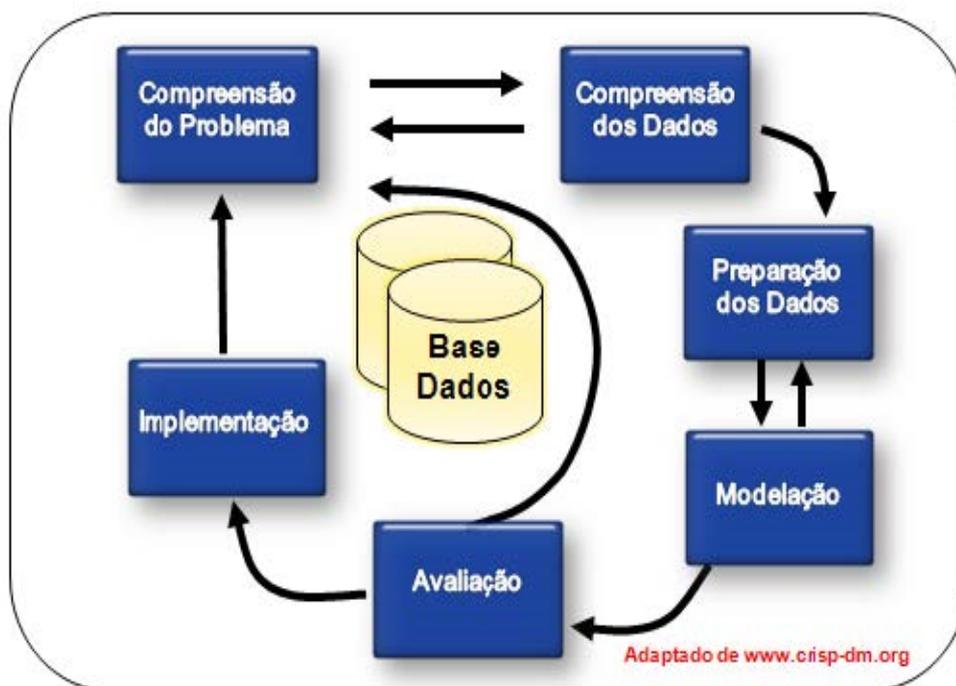
Capítulo 4- Metodologia

4.1 – CRISP – DM

Foram utilizados os softwares *IBM SPSS Modeler 16*, e *IBM SPSS Statistic 20*⁶. A metodologia de DM utilizada pelo *IBM SPSS Modeler* é a CRISP-DM (*Cross Industry Standard Process for Data Mining*), que assenta nas seis fases que estão patentes na figura seguinte e que permitem sistematizar todo o processo de construção do modelo.

De acordo com Adriaans e Zantinge (1996, prefácio - p. v), DM é “a atividade de descobrir conhecimento oculto, padrões inesperados e regras em grandes bases de dados”

Imagem 5– CRISP - DM⁷



Concretamente neste estudo, delimitamos estas várias fases, a saber:

Compreensão do problema: Segundo esta metodologia, a fase inicial de um projeto de DM é a compreensão do problema. Estabelecem-se objetivos,

⁶ IBM SPSS Statistics 20 Algorithms (2011).

⁷ SPSS CRISP-DM 1.0 (2000).

avaliam-se possibilidades, define-se o problema numa perspetiva de DM e traça-se um plano de trabalho inicial.

Nesta dissertação, esta fase compreendeu o estudo da problemática da situação contributiva e fiscal. Avaliaram-se vários modelos de DM possíveis e optou-se pelas RN. Estudaram-se as bases de dados da Informa D&B em busca de informação potencialmente relevante para identificar o perfil das empresas que devem à SS e à AT.

Compreensão dos dados: A fase de compreensão dos dados começou com a compilação das variáveis consideradas de interesse. De seguida procedeu-se à identificação de problemas de qualidade (erros de registo, *missings*, e incoerências).

Preparação dos dados: Esta fase compreende a construção da base de dados final, utilizada posteriormente para alimentar o modelo de DM. Nesta fase, selecionaram-se as variáveis a incluir, fez-se a limpeza de dados (*missings*, *outliers*), a integração de dados (cruzamento de tabelas e agregação de dados por contribuinte) e a construção das variáveis calculadas.

Modelação: Ainda que durante a compreensão do negócio se tenha optado por utilizar RN, nesta etapa decidiram-se com mais detalhe as técnicas a utilizar, tais como: o algoritmo BP, a tangente hiperbólica como função de ativação e o método do GD para minimizar o erro quadrático. Esta escolha baseou-se na pesquisa bibliográfica efetuada e na análise comparativa dos resultados obtidos empregando diferentes técnicas de modelação.

Avaliação: Após a construção do modelo avaliamos a precisão e capacidade de generalização do modelo. Este passo permite avaliar se o modelo desenvolvido pode alcançar os objetivos propostos e identificar eventuais problemas que afetem a qualidade das previsões.

Implementação: Esta fase prevê a implementação prática do modelo, que de momento está fora do âmbito desta dissertação.

É ainda de salientar que a sequência das fases descritas não foi rígida. O CRISP-DM prevê a possibilidade de voltar atrás e refazer parte do procedimento, já que os resultados do modelo dependem em grande parte deste processo de correção e aperfeiçoamento contínuo. Na prática houve

várias situações em que se revelou crucial voltar atrás e repensar as questões previamente analisadas.

4.2 – População, amostra e variáveis

Como referido anteriormente, esta tese tem como objetivo, o desenvolvimento de um modelo matemático que permita identificar, para cada empresa, o risco de ser incumpridor na situação contributiva e tributária.

Para tal utilizou-se um modelo de classificação que, através de RN, atribui a cada empresa, uma pontuação de 1 ou 0, consoante é considerado “contribuinte de risco” ou não, respetivamente.

A designação de “contribuinte de risco” advém da existência de dívidas para com a SS, a AT, ou ambas.

Para a construção do modelo foram consideradas 1.000 empresas, obtidas através da Informa D&B, compreendendo o espaço temporal entre 2011 e 2015. Pretendeu-se focar a investigação na deteção de empresas que possam ter a sua situação contributiva e fiscal em dívida.

A população inicial reportava a todo o sector empresarial de Portugal. Em conjunto com a Informa D&B, seleccionamos uma amostra de 1000 empresas e as variáveis potencialmente relevantes. Foi efetuado o balanceamento dos dados para que cada grupo representasse aproximadamente 50% dos contribuintes da amostra, o que é fulcral em problemas de classificação (Vladislavleva et al., 2010).

Inicialmente foram seleccionadas, como variáveis de entrada, 37 campos da demonstração de resultados e do balanço e 17 indicadores da actividade empresarial⁸.

Após vários ensaios concluímos que usar múltiplos anos não trazia qualquer vantagem para o modelo. A informação tornava-se redundante e as percentagens, nos grupos de teste e de treino, eram muito semelhantes às obtidas utilizando apenas inputs de um ano de actividade.

⁸ Anexo I.

Assim, pareceu-nos mais interessante preservar a simplicidade do modelo, evitando possíveis problemas de *overfitting*. Desta forma, decidimos focalizar a análise no ano mais recente, o de 2015.

Por fim, ainda na ótica de redução do número de variáveis envolvidas, observamos a importância relativa que o modelo atribui a cada variável disponível para o ano selecionado. Optamos por manter apenas as que apresentavam interesse relativo superior a 50%. De seguida, são apresentadas as variáveis selecionadas para o modelo final, elencadas por ordem de importância:

- Resultado_antes_de_depreciações_gasto_e_impostos;
- Participações_financeiras__outros_métodos_2015;
- Estado_e_outros_entes_públicos_2015;
- Resultado_operacional_2015;
- Ano_de_Constituição;
- Total_do_capital_próprio_2015;
- Resultado_antes_de_impostos_2015;
- Outros_rendimentos_e_ganhos_2015;
- Total_do_ativo_corrente_2015;
- Valor_Ações_em_Aberto_Réu;
- Ativos_fixos_tangíveis_2015;
- Outras_variações_no_capital_próprio_2015;
- Ativos_biológicos_Não_Corrente_2015;
- Valor_Ações_Finalizadas_Réu;
- Ativos_intangíveis_2015;
- Diferimentos_2015;
- Valor_Ações_em_Aberto_Autor;
- Ativos_biológicos_Corrente_2015;
- Inventários_2015;
- Total_do_passivo_corrente_2015;
- Ativos_por_impostos_diferidos_2015;
- Vendas_e_serviços_prestados_2015;
- Outros_gastos_e_perdas_2015;

- Financiamentos_obtidos_Não_Corrente_2015;
- Colaboradores_F_2015;
- Capital_Social;
- Outros_instrumentos_de_capital_próprio_2015;
- Outras_contas_a_receber_2015;
- Gastos_com_pessoal_2015;

A amostra é constituída por 1.000 contribuintes distribuídos de acordo com a tabela seguinte:

Tabela I – Caraterização da amostra final

Risco	Contagem de NIF
0	500
1	500
Total Geral	1000

A amostra foi, posteriormente, dividida em 2 grupos: grupo de treino - 679 empresas (68% da amostra); grupo de teste - 321 empresas (32% da amostra). Ambos os grupos são independentes, sendo o grupo de treino necessário à construção do modelo, e o grupo de teste utilizado para medir a capacidade/desempenho de generalização da RN (Zhang, G., 1998).

Tabela II – Amostra de treino e amostra de teste

		Número	Percentagem
Amostra	Treino	679	67,9%
	Teste	321	32,1%
Validos		1000	100,0%
Excluídos		0	
Total		1000	

O método de RN foi aplicado aos registos do grupo de teste, contendo informação sobre as variáveis explicativas e a variável alvo (classificação de risco da situação contributiva e fiscal). O segundo grupo foi reservado para testar a qualidade do modelo.

Capítulo 5- Análise e Discussão de Resultados

Recapitulando, neste trabalho foi utilizada uma RN do tipo MLP, recorrendo ao algoritmo do BP.

Normalmente, a escolha das camadas escondidas do modelo, é uma das questões mais interessantes em termos teóricos. No entanto tal desenvolvimento sai fora do âmbito de análise desta dissertação, remetendo para o Teorema de Aproximação Universal e Séries de Fourier, por isso mesmo não será alvo de aprofundamento. Apesar do *IBM SPSS Modeler* propor uma escolha automática das mesmas (aliás, é possível de todos os parâmetros do modelo), optamos por seleccionar a RN MLP mais reduzida, a de 2 camadas (uma só camada escondida), uma vez que pretendemos apresentar um modelo simples, que de acordo com a literatura existente é computacionalmente suficiente e satisfatório face aos objetivos propostos (Haykin, 2009, p. 167).

Na Tabela V, está definida a nossa RN MLP que delimitamos neste estudo:

Tabela III – Informação RN MLP

Camada de Entrada	Número de nós ^a	29
Camada Escondida	Número de Camadas Escondidas	1
	Número de nós na 1ª Camada Escondida ^a	10
	Função de Ativação	Tangente Hiperbólica
Camada de Saída	Variáveis Dependentes	1
	Número de nós	2
	Função de Ativação	Tangente Hiperbólica
	Função de Erro	Soma dos Quadrados

a. Excluindo a viés

Na tabela IV está o sumário do modelo, que apresenta uma percentagem baixa de previsões incorretas, sendo de 7,8% na fase de teste e de 6% da fase de treino. É indicado que o motivo de paragem do treino é o facto de o erro não diminuir depois de um passo consecutivo do algoritmo, o que pressupõe uma optimização do mesmo.

Concomitantemente, e como não poderia deixar de ser, a soma dos quadrados dos erros vai ao encontro da diminuição da percentagem de previsões incorretas no grupo de teste. Desta forma, a soma dos quadrados dos erros diminui acentuadamente neste segundo grupo (menos de metade do valor calculado para o grupo de treino).

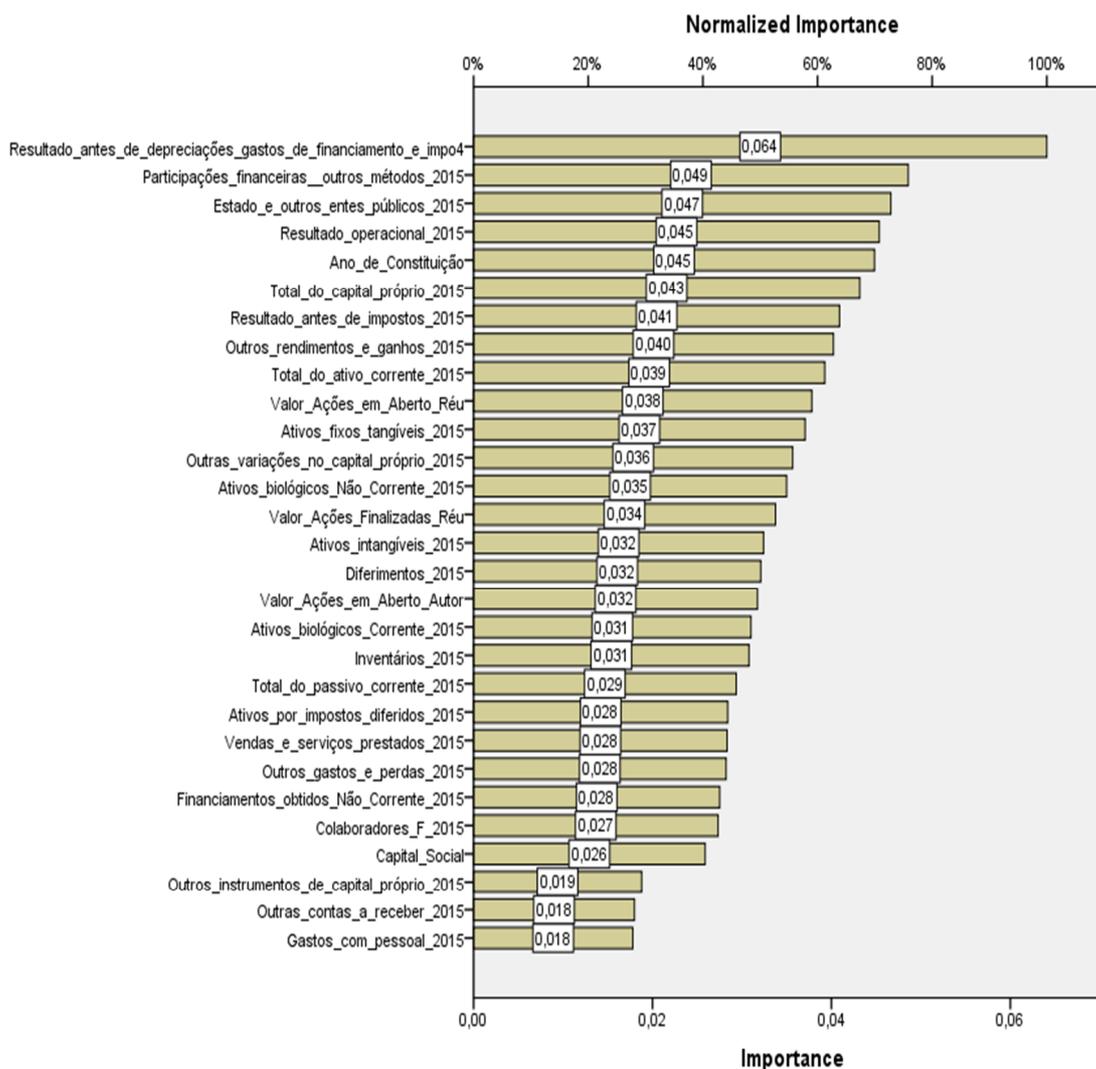
Tabela IV – Resumo do modelo

Treino	Soma dos Quadrados dos Erros	39,387
	Percentagem de Previsões Incorretas	6,0%
	Regra de Paragem Usada	1 passo consecutivo sem diminuição no erro
	Tempo de Treino	0:00:02,07
Teste	Soma dos Quadrados dos Erros	24,804
	Percentagem de Previsões Incorretas	7,8%

Variável Dependente: Risco

A imagem 6 “Importância Normalizada” apresenta a importância relativa das variáveis independentes face à variável independente de maior importância para o modelo, que assume o valor de 100%.

Imagem 6– Importância normalizada das variáveis independentes



No anexo II consta uma tabela com os valores expressos no gráfico anterior. Da análise da imagem 6 verifica-se que a variável de maior importância é o Resultado_antes_de_depreciações_gastos_de_financiamento_e_impostos_2015, sendo seguido por Participações_financeiras_outros_métodos_2015, Estado_e_outros_entes_públicos_2015, Resultado_operacional_2015, e, Ano de Constituição.

Na Tabela V temos os resultados da classificação da RN MLP⁹. No grupo de treino a percentagem total de acerto situa-se nos 94%, sendo de salientar que nos contribuintes considerados sem risco a eficácia do modelo é de 92,7% e nos contribuintes com risco é de 95,3%.

O grupo de teste manteve a excelência nos valores encontrados, validando os resultados anteriores, uma vez que a percentagem total de acerto se situa nos 92,2%, com enfoque nos contribuintes sem risco, onde a percentagem de acerto se situou nos 93%. Nos contribuintes com risco o desempenho do modelo no grupo de teste foi inferior ao grupo de treino, embora com uma excelente percentagem, 91,4%.

A percentagem de identificação global do modelo, tendo em conta os dados consolidados do grupo de treino e do grupo de teste, é de 93,4%.

Tabela V – Classificação da RN MLP

Amostra	Observados	Previstos		
		0	1	Percentagem Acerto
Treino	0	317	25	92,7%
	1	16	321	95,3%
	Percentagem Total	49,0%	51,0%	94,0%
Teste	0	147	11	93,0%
	1	14	149	91,4%
	Percentagem Total	50,2%	49,8%	92,2%

Variável Dependente: Risco

⁹ No anexo III temos o desenho da RN, assim como os resultados finais do seu cálculo.

Capítulo 6- Conclusões

Como referido anteriormente, as empresas possuem um comportamento que se adapta ao período temporal e às suas vicissitudes e fatores contextuais. Estar no meio de uma crise internacional e nacional é a altura oportuna para se estudar qual o comportamento das empresas perante as suas obrigações legais com o Estado português, mais concretamente com as suas contribuições para a SS e o pagamento de impostos com a AT. Tendencialmente as empresas com melhor posição no mercado serão menos avessas a incumprimentos. No entanto, cada qual opta pela sua estratégia de mercado, financeira e fiscal, o que torna fascinante este estudo.

É neste contexto de incerteza quanto ao presente e ao futuro do tecido empresarial e da economia, que se torna evidente a necessidade de modelos robustos de previsão de risco que permitam, por um lado, identificar que tipo de empresas são mais atreitas a não pagar taxas e impostos, e por outro, que estrutura contabilística suporta essa falha. Assim, tal como referido anteriormente, o objetivo fulcral desta tese foi a construção de um modelo de DM que permitisse identificar à priori que empresas terão incumprimento na sua normal e correta situação contributiva e tributária.

Este estudo permitiu retirar algumas conclusões pertinentes. Em primeiro lugar, comprova-se que a partir dos dados existentes é possível obter um modelo com qualidade, alcançando o objetivo proposto. A percentagem de identificação global do modelo, 93,4%, é quase óptima. Embora não exista aqui um termo comparativo com outras abordagens como a regressão linear, os modelos *probit* e *logit*, e a análise multivariada, identificar corretamente mais de 80% de casos como sendo, ou não, de risco, é sinal de que o objetivo inicialmente proposto se concretizou.

De salientar, da análise dos resultados do grupo de teste, que se classificaram erradamente como sendo empresas de risco 11 casos que não têm risco associado, o que corresponde a 7% das empresas nestas condições; do mesmo modo, classificaram-se 14 casos como sendo empresas sem risco mas que na verdade se provou serem de risco, no total de 8,6%. O uso das RN MLP

foi uma boa opção em termos teórico-práticos. Embora raramente usadas no nosso país, provou-se que têm as mesmas têm um grande potencial de aplicação. No caso em apreço, a simplicidade da RN (com 2 camadas, uma escondida), e a escolha das variáveis fizeram a diferença.

Ficou demonstrado que com recurso maioritário a variáveis extraídas da demonstração de resultados e do balanço, é possível obter um modelo robusto e satisfatório. Possivelmente, existirão outras variáveis que podem ajudar a aperfeiçoar o modelo. Este é previsivelmente o passo seguinte: identificar dentro do universo de dados existentes sobre as empresas, quais as que nos podem ser uteis para concluir este puzzle complexo. Noutra vertente, seria interessante utilizar os mesmos dados e verificar comparativamente os resultados obtidos com outras técnicas de análise estatística, e a sua eficácia, assim como, comparar com os indicadores económicos e financeiros.

Contudo, subsistem críticas e limitações ao uso das RN, que apontam para o facto de as mesmas serem uma espécie de caixa negra, deixando-se ao livre arbítrio a conjugação dos dados e a associação de padrões. Outros referem que os resultados excessivamente bons não são mais que distorções do treino supervisionado e da viciação do mesmo. Neste trabalho, a presença de um grupo de teste com resultados semelhantes ao de treino, é deveras encorajador, uma vez que pressupõe que o modelo é robusto e não sofre do efeito atrás descrito, designado por *overfitting* ou sobre ajuste¹⁰. As RN têm defeitos, mas as suas virtudes superam-nas. Grandes quantidades de dados, tratados de formas não convencionais, podem levar ao aparecimento de associações/padrões totalmente inesperados mas verdadeiros. Além do mais, ao longo das últimas 3 décadas e meia, já foram efetuados múltiplos estudos que demonstram a qualidade das RN e provam a teoria matemática subjacente. Por último, esperamos que esta dissertação seja um contributo no sentido de alertar e até apoiar o Estado a construir uma base de dados global para uso próprio e dos interessados, melhorando assim o estudo das empresas e da economia, constituindo-se, também, num instrumento para o desenho de políticas internas e externas, promotoras de desenvolvimento no nosso País.

¹⁰ Ver Capítulo 4 – 4.2 População, amostra e variáveis

Bibliografia

Azevedo, D. (2014), A evolução da Segurança Social em Portugal e o seu enquadramento internacional, Dissertação de Mestrado em Economia, Escola de Economia e Gestão, Universidade do Minho.

Baidyk, T., Kussul, E., & Wunsch, D. (2010), *Neural networks and micromechanics*, Springer, Heidelberg.

Bar, M. (2014), *The Computational Intelligence Techniques for Predictions – Artificial Neural Networks*, em *Annals of the University of Craiova, Economic Sciences Series*, Craiova, Vol. 2, 42, pág.184-190.

Bench-Capon, T.(1993), *Neural networks and open texture*, em *Proceedings of the 4th international conference on Artificial intelligence and law*, Amsterdam, pág. 292-297.

Bishop, C. (2006), *Pattern recognition and Machine Learning*, Springer, New York, Cap.1,5 (pág.1-58; 225-290).

Carpenter, G. & Grossberg, S. (2014), *Adaptive Resonance Theory*, em : Sammut, C. & Webb, G., *Encyclopedia of Machine Learning and Data Mining*, Springer – Verlag, Berlin, pág.1-18.

Chandler, A. D. (1992), *What is a firm?: A historical perspective. European Economic Review*, Elsevier, Vol. 36(2-3), pág.483-492.

Chauvin, Y., & Rumelhart, D. (1995), *Backpropagation: theory, architectures, and applications*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neurais

Chen, J., Su, M., Chen, C., Hsu, F., & Wu, C. (2011), *Application of neural networks for detecting erroneous tax reports from construction companies*, em *Automation in Construction*, Elsevier, Amsterdam, Vol. 20(7), pág.935-939.

Cushing Jr, W., & Arguea, N. (1997), *Neural Network Analysis of the Employee Classification Problem for Tax Purposes*, *Documentos de Trabajo (ICAE)*, (1), pág.1-28.

Cút, S. (2015), *Risk assessment of VAT entities using selected data mining models*, em *Ekonomická revue – Central European Review of Economic*, 18, pág.6-17.

Deng, L., & Yu, D. (2014), *Deep Learning, Foundations and Trends in Signal Processing*, Now Publishers Inc., Hanover - MA, Vol. 7, 3-4, pág.197-387.

Denton, J., Sayeed, L., Perkins, N., & Moorman, A. (1995), *Neural networks to classify employees for tax purposes*, em *Accounting, Management and Information Technologies*, Elsevier, Amsterdam, Vol. 5(2), pág.123-138.

González, P., & Velásquez, J. (2013), *Characterization and detection of taxpayers with false invoices using data mining techniques*, em *Expert Systems with Applications*, Elsevier, Amsterdam, Vol. 40(5), pág.1427-1436.

Gupta, J., & Smith, K. (2000), *Neural networks in business: techniques and applications for the operations researcher*, em *Computers & Operations Research*, Elsevier, Amsterdam, Vol. 27(11), pág.1023-1044.

Haykin, S. (2009), *Neural networks and learning machines*, 3ª Edição, Pearson Education, Prentice Hall, New Jersey, preface – cap.6, pág.1-312.

Hebb, D. (1949), *The Organization of Behavior: A Neuropsychological Approach*, John Wiley & Sons, New York, cap.1-2, pág.1-87.

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neuronais

Hong, S., & Hong, Y. (2007), *Research on Risk Early-Warning System of Social Security Fund Investment*, em *Management Science and Engineering*, ICMSE 2007-International Conference, Harbin, pág. 2040-2045.

Hopfield, J. (1982), *Neural networks and physical systems with emergent collective computational abilities*, em *Proceedings of the national academy of sciences*, National Academy of Sciences, Washington, Vol. 79, pág.2554-2558.

IBM SPSS Statistics 20 Algorithms (2011), IBM, acesso em 31-07-2016, em ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Statistics_Algorithms.pdf.

Kohonen, T. (1982), *Self-organized formation of topologically correct feature maps*, em *Biological cybernetics- Advances in Computational Neuroscience*, Springer-Verlag, Berlin, Vol. 43, pág.59-69

Leskovec, J., Rajaraman, A., & Ullman, J. (2014), *Mining of Massive Datasets*, Stanford University, acesso em 20-08-2016, disponível em <http://infolab.stanford.edu/~ullman/mmds/book.pdf>.

McCulloch, S., & Pitts, W. (1943), *A logical calculus of the ideas immanent in nervous activity*, em *The bulletin of mathematical biophysics*-University of Oxford, Pergamon Press, Oxford, Vol. 5, pág.115-133.

Minsky, M. & Papert, S. (1969), *Perceptrons : An Introduction to Computational Geometry*, MIT Press, Cambridge – MA.

Norregaard, J., & Khan, T. S. (2007), *Tax policy: Recent trends and coming challenges*, No. 7-274, *International Monetary Fund Working Paper*, Washington.

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neuronais

Paliwal, M., & Kumar, U. (2009), *Neural networks and statistical techniques: A review of applications*, em *Expert systems with applications*, Elsevier, Vol. 36.1, pág.2-17.

Páscoa, A. (2009), Enquadramento de legislação nos Sistemas de Segurança Social, Dissertação de Mestrado em Contabilidade - ramo Fiscalidade, ISCA, Universidade de Aveiro.

Pereira, M.F. (2011), Fiscalidade, 4ª Edição, Almedina, Coimbra, pág.113-120.

Pinto, M. P. F. (2014), A Crise de Dívida Soberana na Área do Euro, Dissertação de Mestrado em Economia, Faculdade de Economia do Porto.

Rojas, R. (1996), *Neural networks: a systematic introduction*, Springer-Verlag, Berlin, cap.1-8, pág.3-227.

Shaoqiu, Zhang & Yingying, Wei (2010), *Research on Tax Prediction Model Based on BP Neural Network*, paper, disponível em <http://www.seiofbluemountain.com/upload/product/201001/1263265486s7rptjk6.pdf>, acesso em 10-09-2016.

Silva, S. (2015), Direito fiscal: teoria geral, 2ª Edição, Imprensa da Universidade de Coimbra, Coimbra, pág.32-51.

SPSS CRISP-DM 1.0 (2000), SPSS, disponível em <https://www.the-modeling-agency.com/crisp-dm.pdf>, acesso em 1-09-2016.

Tkáč, M., & Verner, R. (2016), *Artificial neural networks in business: Two decades of research*, em *Applied Soft Computing*, Elsevier, Vol. 38, pág.788-804.

Bruno Mendes

Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes neuronais

Vladislavleva, E., Smits, G., & Den Hertog, D. (2010), *On the importance of data balancing for symbolic regression*, em *IEEE Transactions on Evolutionary Computation*, New York, Vol.14(2), pág.252-277.

Wong, B., & Selvi, Y. (1998), *Neural network applications in finance: A review and analysis of literature (1990–1996)*, em *Information & Management*, Elsevier, Vol. 34(3), pág.129-139.

Wong, B., Bogdanovich, T., & Selvi, Y. (1997), *Neural network applications in business: A review and analysis of the literature (1988–1995)*, em *Decision Support Systems*, Elsevier, Amsterdam, Vol.19, pág.301-320.

Wong, B., Lai, V., & Lam, J. (2000), *A bibliography of neural network business applications research: 1994–1998*, em *Computers & Operations Research*, Elsevier, Amsterdam, Vol. 27(11), pág.1045-1076.

Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998), *Forecasting with artificial neural networks: The state of the art*, em *International journal of forecasting*, Elsevier, Amsterdam, Vol.14(1), pág.35-62.

ANEXOS

ANEXO I – Variáveis independentes iniciais**Tabela VI – Variáveis independentes iniciais***

Nome	Caraterização	Descritivo
Capital Social	Base de Dados (BD)	Montante de entrada para o início da actividade
Ano de Constituição	BD	Ano de registo da empresa na conservatória registo comercial / AT
NºSócios/Acionistas	BD	Constantes na conservatória registo comercial / AT
Nº Órgãos Sociais	BD	Administradores, Conselho de Administração, Comissão Executiva, Gerente, Sócio, Sócio-Gerente
Certificações	BD	Normas Internacionais e Portuguesas (NP, NP ISO)
Distinção PME Líder/Excelência	BD	Prémios ganhos
NºAções em Aberto (Autor)	BD	Número de processos pendentes em Tribunal
Valor Ações em Aberto (Autor)	BD	Valor de processos pendentes em Tribunal
NºAções Finalizadas (Autor)	BD	Número de processos em Tribunal concluídos
Valor Ações Finalizadas (Autor)	BD	Valor de processos em Tribunal concluídos
Nº Ações em Aberto (Réu)	BD	Número de processos pendentes em Tribunal, como réu
Valor Ações em Aberto (Réu)	BD	Valor de processos pendentes em Tribunal, como réu
NºAções Finalizadas (Réu)	BD	Número de processos em Tribunal concluídos, como réu
Valor Ações Finalizadas (Réu)	BD	Valor de processos em Tribunal concluídos, como réu
Colaboradores 2011	BD	Número total de funcionários da empresa
Colaboradores (M) 2011	BD	Número de funcionários da empresa, sexo masculino
Colaboradores (F) 2011	BD	Número de funcionários da empresa, sexo feminino
Ativos fixos tangíveis 2011	Balço (B)	Itens tangíveis que sejam detidos para uso na produção ou fornecimento de bens ou serviços, para arrendamento a outros, ou para fins administrativos; e se espera que sejam usados durante mais do que um período
Ativos intangíveis 2011	B	Ativo é um recurso controlado por uma entidade como resultado de acontecimentos passados; e do qual se espera que fluam benefícios económicos futuros para a entidade. Ativo intangível: é um activo não monetário identificável sem substância física
Ativos biológicos (Não Corrente) 2011	B	Ativo biológico é um animal ou planta vivos, que não seja considerado como corrente

Nome	Caraterização	Descritivo
Ativos por impostos diferidos 2011	B	Quantias de impostos sobre o rendimento recuperáveis em períodos futuros, respeitantes a: -Diferenças Temporárias Dedutíveis;- Reporte de perdas fiscais não utilizadas;- Reporte de créditos tributáveis não utilizados
Participações financeiras - outros métodos 2011	B	Quantias de investimentos e interesses, em subsidiárias, em associadas e em entidades conjuntamente controladas, para as quais, no balanço individual ou no balanço consolidado, se adopte outro método, que não seja o método da equivalência patrimonial
Total do ativo não corrente 2011	B	Todos os outros ativos não classificados como correntes
Inventários 2011	B	Ativos detidos para venda no decurso ordinário da actividade empresarial; no processo de produção para tal venda; ou na forma de materiais ou consumíveis a serem aplicados no processo de produção ou na prestação de serviços
Ativos biológicos (Corrente) 2011	B	Ativo biológico é um animal ou planta vivos.
Clientes 2011	B	Compradores de mercadorias, de produtos e de serviços
Outras contas a receber 2011	B	Quantias de ativos financeiros correspondentes a contas a receber que não estejam inseridas nas demais rubricas de contas a receber, e cujo tratamento é estabelecido na NCRF 27 – Instrumentos Financeiros
Diferimentos 2011	B	Todos os custos e proveitos que devam ser reconhecidos nos exercícios seguintes
Caixa e depósitos bancários 2011	B	Inclui meios líquidos de pagamento, e, meios financeiros disponíveis em instituições financeiras
Total do ativo corrente 2011	B	O conjunto dos ativos que satisfaçam qualquer dos seguintes critérios: Espera-se que seja realizado, ou pretende-se que seja vendido ou consumido, no decurso normal do ciclo operacional da entidade; Esteja detido essencialmente para a finalidade de ser negociado; Seja realizado num período até doze meses após a data do balanço; ou seja caixa ou equivalente de caixa, a menos que lhe seja limitada a troca ou sejam usados para liquidar um passivo durante pelo menos doze meses após a data do balanço
Capital realizado 2011	B	Corresponde à diferença entre a quantia de capital estatutário da entidade e a parte não realizada pelos sócios
Outros instrumentos de capital próprio 2011	B	Quantias relativas a outros instrumentos classificáveis como de capital próprio, desde que obedeam à definição de instrumentos de capital próprio da NCRF 27 – Instrumentos Financeiros
Reservas legais 2011	B	De acordo com o Código das Sociedades Comerciais é obrigatória a constituição de uma reserva legal. São lucros que não podem ser distribuídos
Resultados transitados 2011	B	Regista os resultados líquidos e os dividendos antecipados, provenientes do exercício anterior
Outras variações no capital próprio 2011	B	Compreende os subsídios ao investimento
Total do capital próprio 2011	B	Corresponde ao património líquido da empresa, e é grosso modo igual ao ativo menos o passivo duma empresa
Provisões 2011	B	Constituídas pelos passivos relativamente aos quais a sua quantia ou tempestividade são incertos
Financiamentos obtidos (Não Corrente) 2011	B	São os financiamentos obtidos, sejam eles de instituições de crédito e sociedades financeiras, ou de outras entidades, não considerados como correntes
Total do passivo não corrente 2011	B	Obrigação presente que seja fruto de acontecimentos do passado e da qual se espera um exfluxo, ou uma saída de fundos da entidade, que não seja corrente

Nome	Caraterização	Descritivo
Fornecedores 2011	B	Vendedores de bens e de serviços, com exceção dos destinados ao imobilizado
Estado e outros entes públicos 2011	B	Relações com o Estado, autarquias locais e outros entes públicos que tenham características de impostos e taxas
Financiamentos obtidos (Corrente) 2011	B	São os financiamentos obtidos, sejam eles de instituições de crédito e sociedades financeiras, ou de outras entidades
Outras contas a pagar 2011	B	Categoria residual, na qual são registados passivos não correntes sem enquadramento nas restantes rubricas
Total do passivo corrente 2011	B	Quando satisfizer qualquer dos seguintes critérios: espera-se que seja liquidado durante o ciclo operacional normal da entidade; está detido essencialmente para a finalidade de ser negociado; deverá ser liquidado num período até doze meses após a data do balanço; ou a entidade não tem um direito incondicional de diferir a liquidação do passivo durante pelo menos doze meses após a data do balanço.
Vendas e serviços prestados 2011	Demonstração de resultados (DR)	Vendas, no caso de produtos tangíveis ou dos serviços prestados, quando o produto é intangível, deduzidos do IVA ou outros impostos indiretos sobre o consumo
Custos das mercadorias vendidas e das matérias	DR	Gastos diretamente relacionados com a produção dos bens e serviços vendidos
Fornecimentos e serviços externos 2011	DR	Genericamente é a aquisição de um conjunto de bens e serviços, de consumo imediato, que a generalidade das empresas necessitam para o seu funcionamento normal
Gastos com pessoal 2011	DR	Salários, encargos para a Segurança Social a cargo da empresa, despesas com refeições dos colaboradores, prémios e outros encargos diretamente relacionados com a força de trabalho
Outros rendimentos e ganhos 2011	DR	Rendimentos e ganhos não classificados nas rubricas da DR
Outros gastos e perdas 2011	DR	Gastos e perdas não classificados nas rubricas da DR
Resultado antes de depreciações, gastos de financiamento e impostos 2011	DR	EBITDA - ao eliminar os efeitos das decisões de financiamento, permite analisar e comparar o desempenho/rentabilidade entre empresas
Resultado operacional 2011	DR	EBIT - permite comparações sem ter em conta os efeitos derivados de diferentes estruturas de capital e taxas de imposto; Igual a Proveitos Operacionais – Custos Operacionais
Resultado antes de impostos 2011	DR	EBT - permite estabelecer comparações entre entidades sujeitas a diferentes jurisdições fiscais
Resultado líquido do período 2011	DR	Lucro que a empresa apresenta num dado período

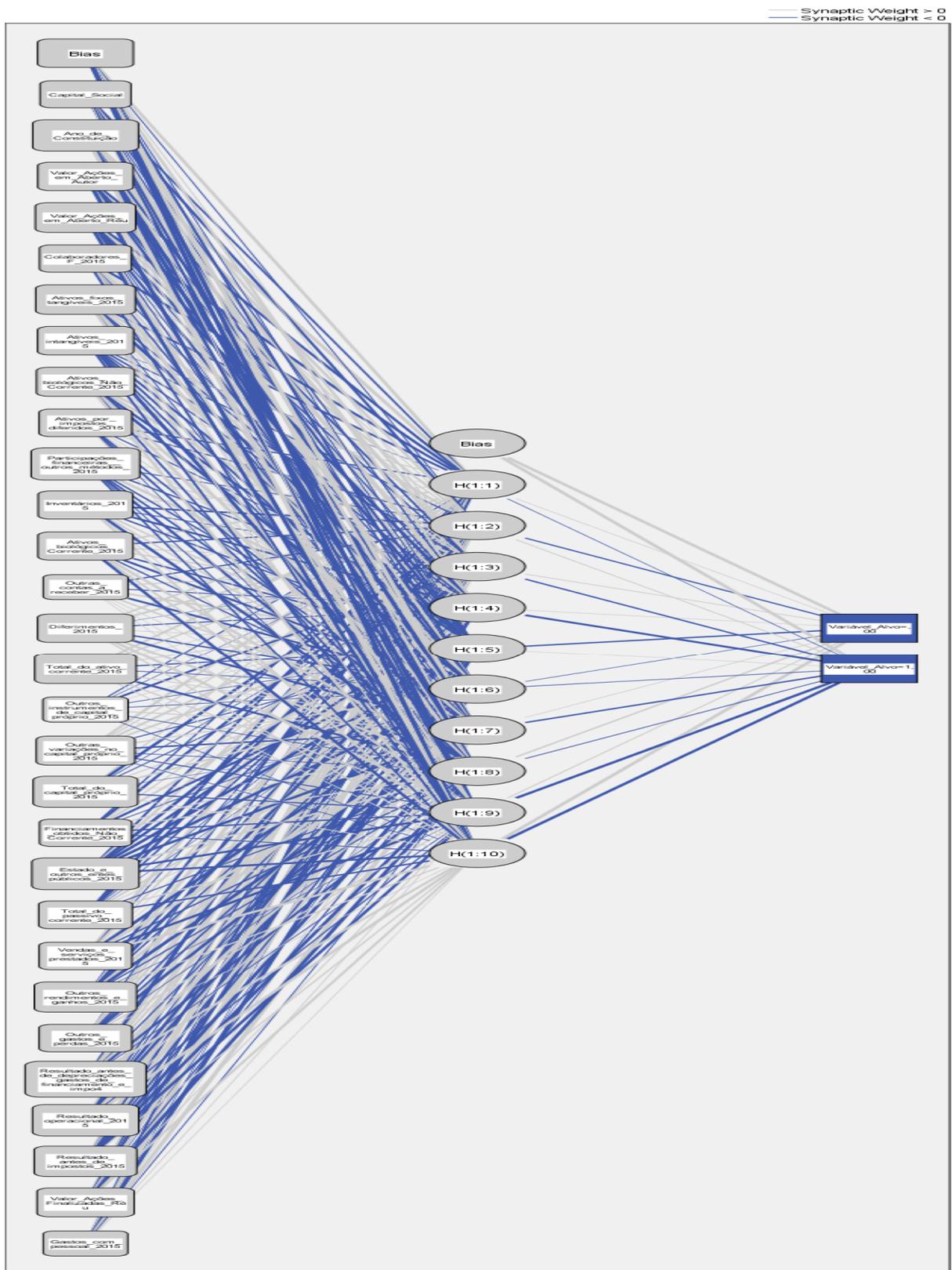
* Balanço e Demonstração de resultados com dados de 2011, 2012, 2013, 2014, 2015. Descrições efetuadas de acordo com o SNC e NCRF em vigor.

ANEXO II – Importância das variáveis independentes**Tabela VII – Importância das variáveis independentes**

Variáveis Independentes	Importância	Importância Normalizada
Capital_Social	,026	40,4%
Ano_de_Constituição	,045	69,9%
Valor_Ações_em_Aberto_Autor	,032	49,5%
Valor_Ações_em_Aberto_Réu	,038	59,0%
Colaboradores_F_2015	,027	42,6%
Ativos_fixos_tangíveis_2015	,037	57,8%
Ativos_intangíveis_2015	,032	50,6%
Ativos_biológicos_Não_Corrente_2015	,035	54,6%
Ativos_por_impostos_diferidos_2015	,028	44,3%
Participações_financeiras_outros_métodos_2015	,049	75,8%
Inventários_2015	,031	48,1%
Ativos_biológicos_Corrente_2015	,031	48,4%
Outras_contas_a_receber_2015	,018	28,1%
Diferimentos_2015	,032	50,1%
Total_do_ativo_corrente_2015	,039	61,2%
Outros_instrumentos_de_capital_próprio_2015	,019	29,3%
Outras_variações_no_capital_próprio_2015	,036	55,7%
Total_do_capital_próprio_2015	,043	67,4%
Financiamentos_obtidos_Não_Corrente_2015	,028	42,9%
Estado_e_outros_entes_públicos_2015	,047	72,8%
Total_do_passivo_corrente_2015	,029	45,8%
Vendas_e_serviços_prestados_2015	,028	44,2%
Outros_rendimentos_e_ganhos_2015	,040	62,8%
Outros_gastos_e_perdas_2015	,028	44,0%
Resultado_antes_de_depreciações_gastos_de_financiamento_e_impo4	,064	100,0%
Resultado_operacional_2015	,045	70,8%
Resultado_antes_de_impostos_2015	,041	63,8%
Valor_Ações_Finalizadas_Réu	,034	52,7%
Gastos_com_pessoal_2015	,018	27,8%

ANEXO III – Desenho e cálculos da RN MLP

Imagem 7 – Desenho da RN MLP



Hidden layer activation function: Hyperbolic tangent
Output layer activation function: Hyperbolic tangent

Tabela VIII - Estimativas de Parâmetros

Preditor		Previsto											
		Camada Escondida 1										Camada de Saída	
		H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	H(1:9)	H(1:10)	[corr_Sum=0,00]	[corr_Sum=1,00]
Camada de entrada	(Bias)	,385	-,072	,055	,312	-,069	-,277	1,007	,304	-,667	-1,374		
	Capital_Social	-,219	-,312	,321	-,142	-,072	-,103	,385	-,289	-,284	,082		
	Ano_de_Constituição	,519	,368	1,835	1,238	-,301	,345	-,681	-,664	-,349	,964		
	Valor_Ações_em_Aberto_Autor	-,342	,075	-,237	-,152	-,369	-,332	,224	,161	,019	-,522		
	Valor_Ações_em_Aberto_Réu	-,413	,312	-,185	-,416	-,126	-,300	-,046	-,101	,333	-,431		
	Colaboradores_F_2015	-,368	,114	,219	,296	-,100	,263	-,153	,504	,107	-,232		
	Ativos_fixos_tangíveis_2015	-,339	,424	,245	-,202	-,519	,539	,022	-,063	,292	,134		
	Ativos_intangíveis_2015	-,018	-,175	-,226	-,163	-,292	-,271	-,206	-,528	,021	-,258		
	Ativos_biológicos_Não_Corrente_2015	,138	,084	,491	,411	-,363	-,440	,391	,484	-,219	,356		
	Ativos_por_impostos_diferidos_2015	-,282	-,221	,105	-,147	,046	,056	-,212	-,257	-,042	,029		
	Participações_financeiras_outros_mét_2015	,038	-,152	,184	-,234	,584	-,208	-,241	-,367	-,080	-,209		
	Inventários_2015	,136	-,437	-,124	-,456	-,248	,258	,220	-,466	,050	,307		
	Ativos_biológicos_Corrente_2015	,432	-,334	,179	,306	-,332	-,410	-,184	,486	,270	-,410		
	Outras_contas_a_receber_2015	,041	-,283	,222	-,205	,465	-,476	,309	,045	,036	,038		
	Diferimentos_2015	,362	,487	-,004	-,239	,130	-,320	-,679	,502	,584	,198		
	Total_do_ativo_corrente_2015	,401	,286	,069	,171	-,280	-,131	,176	-,494	,230	,085		
	Outros_instrumentos_de_capital_próprio_2015	-,134	-,016	-,400	,067	-,080	-,126	,255	-,102	,035	-,053		
	Outras_variações_no_capital_próprio_2015	,416	,180	,414	,392	-,169	,279	-,385	-,102	-,198	-,075		
	Total_do_capital_próprio_2015	,173	-,477	-,048	-,126	,161	,366	-,013	,021	-,273	,151		
	Financiamentos_obtidos_Não_Corr_2015	-,160	-,379	,151	,359	-,398	-,098	,246	,300	-,311	-,157		
	Estado_e_outros_entes_públicos_2015	-,285	,361	-,487	-,407	-,688	-,727	-	,376	-,2159	-,3497		
	Total_do_passivo_corrente_2015	-,356	-,356	-,218	-,036	-,009	,295	,085	-,164	,226	-,327		
	Vendas_e_serviços_prestados_2015	,480	-,087	,254	,007	,016	-,569	,541	-,209	-,590	,782		
Outros_rendimentos_e_ganhos_2015	-,346	,416	-,286	-,448	-,188	,645	-,676	-,161	-,013	,208			
Outros_gastos_e_perdas_2015	-,138	,353	,246	,307	-,177	-,207	-,453	-,030	,192	,032			
Resultado_antes_de_dep_gastos_e_impo4	,282	-,512	,286	,293	,260	-,134	,481	-,339	-,064	,353			
Resultado_operacional_2015	,357	,343	-,166	-,094	-,378	-,342	,170	,077	-,134	,462			
Resultado_antes_de_impostos_2015	,362	,424	-,306	,442	-,262	-,195	-,478	-,420	-,380	,146			
Valor_Ações_Finalizadas_Réu	,363	-,147	-,084	-,443	,403	-,007	-,115	-,507	,080	,287			
Gastos_com_pessoal_2015	-,257	,281	,507	-,006	-,719	,036	,914	-,066	,044	,024			
Camada eEscondida 1	(Bias)										1,087	,928	
	H(1:1)										,042	-,036	
	H(1:2)										-,221	,103	
	H(1:3)										,105	-,277	
	H(1:4)										,062	-,589	

Bruno Mendes
 Estudo do incumprimento da situação contributiva e fiscal das empresas utilizando redes
 neuronais

Preditores	Previsto											
	Camada Escondida 1										Camada de Saída	
	H(1:1)	H(1:2)	H(1:3)	H(1:4)	H(1:5)	H(1:6)	H(1:7)	H(1:8)	H(1:9)	H(1:10)	[corr_Sum=0,00]	[corr_Sum=1,00]
H(1:5)											-,313	,039
H(1:6)											-,053	-,094
H(1:7)											,064	-,310
H(1:8)											,257	-,108
H(1:9)											,130	-,780
H(1:10)											1,159	-,621