

## " UMA NOVA METODOLOGIA EM FINANÇAS: AS REDES NEURONAIIS." "A NEW METHODOLOGY IN FINANCE: THE NEURAL NETWORKS"

Elísio Brandão

"Docteur D'État en Gestion: Université de Paris IX-Dauphine"  
e Professor da Universidade do Porto (Faculdade de Economia)

El objetivo de este trabajo es presentar una nueva aproximación a las Finanzas Corporativas. En primer lugar, comenzamos mostrando el método clásico consistente en la percepción lineal y no lineal. Finalmente, presentamos los mapas auto organizados de Kohonen.

The goal of this paper is to present a new approach in Corporate Finance. Firstly, we start by showing the classic method followed by the linear and non-linear perception. Finally we present the self-organised maps of Kohonen.

**PALABRAS CLAVE:** redes neurales de Kohonen, retroalimentación, percepción, método descendente de gradiente estocástico, rango constante del sector

**KEY WORDS:** Kohonen neural network, feed-back, perceptions, backpropagation, stochastic gradient descent method, constant neighbourhood range.

Muitos dos modelos utilizados em finanças baseiam-se em relações lineares entre as variáveis e no pressuposto de que elas seguem uma distribuição normal reduzida. São os casos dos modelos de equilíbrio dos activos financeiros (C.A.P.M.)<sup>145</sup>, do ICAPM e dos modelos de previsão de falência. O caminho aleatório, "random walk", das cotações bolsistas não é ele caótico<sup>146</sup>, isto é, com aparência somente devido ao azar mas na realidade seguindo processos determinísticos complexos? Estes modelos deram lugar à utilização de regressões com funções logarítmicas, semi-logarítmica ou lineares por partes, para procederem à estimação empírica dos parâmetros dos modelos.

Estas técnicas limitaram a profundidade e o rigor da análise devido aos instrumentos estatísticos utilizados. Weigend (1991) ao efectuar o estudo dos fenómenos caóticos, com base na série recorrente (iterated quadratic map) os coeficientes de auto correlação apresentados, para os quatro primeiros atrasos, poderiam sugerir um fenómeno aleatório mas o sistema que o produz é bem determinístico<sup>147</sup>.

**As redes neuronais** conheceram nestes últimos anos, um grande interesse devido ao desenvolvimento dos métodos utilizados, nomeadamente o algoritmo retro-propagação dos erros e de algumas das suas propriedades (a aproximação universal das funções) e que fizeram nascer a esperança de poder tirar partido de relações não lineares existentes nos mercados financeiros e nas empresas.

O estudo das redes neuronais têm interesse nos domínios da gestão financeira na medida em que elas constituem um prolongamento dos métodos econométricos e estatísticos de previsão e de análise dos dados.

Num primeiro tempo apresentaremos os métodos clássicos seguido do perceptrão linear e não linear, as suas vantagens e desvantagens bem como os modos de utilização.

Por último exporemos as cartas auto-organizadas, em particular a rede estruturada de Kohonen.

<sup>145</sup> Este modelo estabelece uma relação linear entre a rentabilidade e o risco.

<sup>146</sup> Os mercados têm um comportamento linear ou os instrumentos de análise utilizados até ao presente não permitem detectar outro tipo de comportamento, nas interdependências complexas nas séries dos preços.

<sup>147</sup> O sistema é determinístico porque se conhece a equação só que o instrumento estatístico utilizado para captar as interdependências, isto é, os coeficientes de correlação não permitem detectar fenómenos não lineares.

## 1. PERCEPTRÃO “MULTI-NÍVEL”

A título de introdução, admitamos que se pretendia elaborar um modelo de regressão, baseado em três indicadores, o da rentabilidade, liquidez e estrutura financeira, cuja finalidade era a emissão de um parecer, sobre a saúde financeira de uma amostra de empresas.

De acordo com a abordagem clássica o modelo de regressão linear assume a forma seguinte:

$$C_i = \alpha + \beta_1 * ROE_i + \beta_2 * (D/S)_i + \beta_3 * AT_i + \varepsilon_i$$

onde:

$C_i$  representa a avaliação (classificação) da saúde financeira da empresa  $i$ ;

$ROE_i$ ,  $(D/S)_i$ ,  $AT_i$  são a rentabilidade financeira, a estrutura financeira e a liquidez geral da empresa  $i$ , respectivamente;

$\alpha$ ,  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são os coeficientes da regressão aplicável ao perceptrão linear;

$\varepsilon_i$  é o resíduo do modelo.

Os parâmetros seriam estimados pelo método dos mínimos quadrados.<sup>148</sup>

Mas, se o coeficiente de correlação fossem pequenos, os coeficientes da regressão podiam não ser significativamente diferentes de zero, para um intervalo de confiança de 95%, e existir uma relação não linear na fronteira entre as empresas sãs e as empresas em dificuldades financeiras.

## 2. PERCEPTRÃO LINEAR

Este modelo de regressão permite construir uma relação linear entre as variáveis de entrada ( $ROE$ ,  $D/S$  e  $AT$ ) e a variável de saída<sup>149</sup> ( $C_i$ ).

A variável de saída corresponde à soma ponderada das variáveis de entrada à qual se acresce uma ordenada na origem. O modelo de regressão é uma forma de **processador** da informação. Ele recebe a informação, trata-a, através de uma soma ponderada, e reenvia os resultados.

O **operador**  $\Sigma$  representa a soma das variáveis de entrada ponderadas pelos coeficientes  $\beta_1$  e  $\beta_2$  e  $\beta_3$  das variáveis de entrada. Estes coeficientes<sup>150</sup> são designados ponderadores, e caracterizam as conexões entre a célula do operador e as variáveis de entrada<sup>151</sup>.

O coeficiente  $\alpha$  é a ponderação, isto é, o peso que caracteriza a conexão com uma célula de entrada que toma como único valor a unidade<sup>152</sup>.

O resultado obtido é a classificação da empresa,  $C$ .<sup>153</sup>

Com estes procedimentos acabamos de elaborar a primeira **rede de neurónios** que corresponde a um outro modo de reescrever o modelo de regressão inicial.

Trata-se de um perceptrão linear a dois níveis. Um nível de células de entrada composto de quatro células<sup>154</sup> e um nível de células de saída composto de uma única célula<sup>155</sup>.

O número de níveis de células, o número de células por nível e o esquema de conexão entre as células definem a "topologia" da rede<sup>156</sup>.

<sup>148</sup> O algoritmo dos mínimos quadrados é um procedimento de estimação dos parâmetros que minimiza a soma dos quadrados dos desvios entre os valores estimados e os reais calculados com o auxílio do modelo.

<sup>149</sup> A classificação da empresa.

<sup>150</sup>  $\alpha$ ,  $\beta_1$ ,  $\beta_2$  e  $\beta_3$  são os parâmetros do perceptrão e são determinados em função da aprendizagem.

<sup>151</sup> Também designadas células de entrada.

<sup>152</sup> Desempenha as funções de constante na regressão.

<sup>153</sup> Posteriormente, interessar-nos-emos pela "transfer function" ou de activação  $\Psi$ . A este nível de análise só nos interessa o operador identidade.

<sup>154</sup> As três variáveis e o resíduo.

<sup>155</sup> A classificação.

Sejam:

$W_{ij}$  os pesos que caracterizam as conexões entre duas células, sendo  $i$  a célula de origem e  $j$  a célula de destino;

$U_i$  é a soma das entradas de uma célula  $i$  ponderada pelos pesos que as rellam às células do nível precedente (o resultado do operador  $\Sigma$ );

$X_i$  é o resultado da aplicação da função de transferência (transfer function) à  $U_i$  para o nível (couche) de entrada e os níveis intermédios,  $y_i$ , para o nível de saída.

Para estimar os parâmetros deste perceptrão (os coeficientes ou pesos  $W_{ij}$ ) minimiza-se a soma, para todas as empresas, do quadrado dos erros.<sup>157</sup>

Recorre-se não a um método directo tal como a inversão da matriz mas a um método iterativo de ajustamento progressivo designado por método do gradiente.

Isso define o algoritmo de aprendizagem ou algoritmo de estimação.

Como a função erro é quadrada definida e positiva ela só tem um mínimo global.

Procura-se, pois, o mínimo da função erro ou função objectivo:

$$E = \frac{1}{2} \cdot \sum_{p=1}^h \left( d_p - \hat{y}_p \right)^2$$

onde:

o índice  $p$  representa o número de empresas da amostra;

$d_p$  é o valor desejado para a empresa  $p$ ;

$$\hat{y}_p$$

é o valor calculado, isto é,

$$\hat{y}_p = \sum_{k=1}^n w_k x_k$$

$n$  é o número de células do nível precedente.

O **gradiente descendente** é uma abordagem intuitiva que consiste no ajustamento dos parâmetros da função objectivo com a finalidade de reduzir o erro global. O ajustamento é efectuado seguindo a direcção indicada pelo oposto do vector gradiente, isto é, o vector da primeira derivada da função objectivo. Assim, temos que:

$$\Delta W = - \eta \left( \frac{\partial E}{\partial W} \right)$$

onde:

$\eta$  é um coeficiente de aprendizagem que representa a amplitude da deslocação que o algoritmo realiza no sentido da recta .

O que significa que:

$$\Delta W_k = - \eta \left( \frac{\partial E}{\partial W_k} \right) = \eta \sum_p (d_p - \hat{y}_p) x_k$$

<sup>156</sup> A "topologia" da rede define as características da rede.

<sup>157</sup> Mais propriamente o quadrado dos desvios entre o valor desejado e o valor calculado.

Se  $\eta$  é suficientemente pequeno, pode-se mostrar a convergência do algoritmo em direcção ao mínimo que neste caso é único.

### 3. OS PERCEPTRÕES NÃO LINEARES

A hipótese de linearidade limita fortemente o campo de aplicação potencial desta abordagem.

Para dotar o perceptrão de uma capacidade de modelização não linear é, vulgarmente, utilizada uma "transfer function" que modifica a topologia da rede.

Esta função "sigmoidal",  $\Psi$ , é expressa pela equação:

$$\Phi(\mu) = \frac{1}{1 + e^{-\eta\mu}}$$

Se  $\mu$  é a soma das entradas de uma célula ponderada pelos pesos que a reiam às células do nível (couche) precedente, a função  $\Psi$  é diferenciável em relação aos parâmetros da rede<sup>158</sup> e limitada entre 0 e 1.

O parâmetro  $\gamma$  é um coeficiente positivo que define a recta da sigmoide na origem.

As células que utilizam esta "transfer function" são ditas não lineares.

Várias topologias podem ser testadas. Elas caracterizam-se pela introdução de um nível de células ditas escondidas situando-se entre as células de entrada e as células de saída. No caso presente, de **duas** células e pelo carácter linear (L) ou não linear (NL) da célula de saída.

Surge um problema ao nível do algoritmo de aprendizagem. A abordagem do gradiente descendente supõe uma medida directa da diferença entre o valor desejado ( $d_p = (d_{1p} \dots d_{ep})$ ) e o valor calculado  $y_p = (y_{1p} \dots y_{ep})$ . Ora, o valor desejado não é explícito para as células do nível escondido.

Para resolverem o problema acabado de levantar, Rumelhart, Hinton e Williams (1986) propuseram o algoritmo dito de retro-propagação dos erros (error backpropagation).

Trata-se com efeito de uma aplicação do gradiente descendente instantâneo (ou estocástico) baseando-se sobre a regra da derivada em cadeia (ou derivada da função composta).

As fórmulas de adaptação dos pesos são as seguintes:

$$\Delta \omega_{ij} = \eta \delta_j x_i$$

onde:

$\delta_j = x_j \cdot (1 - x_j) \cdot (d_j - y_j)$  e  $j$  representa uma célula de saída não linear.

$\delta_j = 1 \cdot (d_j - y_j)$  e  $j$  representa uma célula de saída linear.

$\delta_j = x_j \cdot (1 - x_j) \cdot \sum_{k=1}^n \delta_k \cdot w_{jk}$ , e  $j$  representa uma célula escondida e  $k$  as células que se lhe seguem.

Existem numerosas variantes quanto às modalidades de aplicação de este algoritmo.<sup>159</sup>

A utilização dos **perceptrões** apresenta algumas **vantagens** mas também **desvantagens**.

No que respeita às vantagens podemos referir as seguintes:

- O nível de erro<sup>160</sup> é normalmente inferior com a ajuda da rede de neurónios não lineares relativamente à abordagem linear;

<sup>158</sup> Os pesos que caracterizam as conexões entre as células.

<sup>159</sup> Cichocki e Umbehnen (1993) fazem uma síntese.

- Quanto maior for a não linearidade da rede mais a repartição do espaço obtido se aproxima da célula sugerida pelos dados;
- Quanto mais a topologia da rede é complexa<sup>161</sup> mais a fronteira obtida é sinuosa ou deformada;
- Por último, refira-se que a introdução do perceptrão "multi-níveis" e o algoritmo de retro-propagação não implicam colocar nenhuma hipótese à priori sobre a forma de distribuição das variáveis observadas, sobre a forma dos resíduos do modelo ou sobre a existência de interacções entre as variáveis.<sup>162</sup>

Quanto às **desvantagens** podemos citar as seguintes:

- A determinação da topologia é um problema que só conhece soluções parciais.<sup>163</sup>
- A convergência da aprendizagem levanta duas questões: o algoritmo converge em direcção a que mínimo da função erro e a que velocidade é que ele converge?

No que se refere à topologia para Takahasi (1993), uma das propriedades estatísticas e matemáticas mais remarcáveis dos perceptrões é serem "aproximadores universais de funções. Contudo, esta propriedade só é verificada sobre um suporte fechado e limitado e que se dote o perceptrão de células suficientes<sup>164</sup> sobre a "couche" escondida.

Quanto ao mínimo e à velocidade da convergência estas duas questões encontram uma resposta clara no contexto linear em que a função erro é quadrada e só existe um mínimo global ao qual conduz todos os algoritmos numéricos de minimização, quer sejam de primeira ou de segunda ordem, deterministas ou estocásticos.

A situação é mais complexa no contexto não linear. A função objectivo ou erro cometido é uma função complexa, que conhece numerosos mínimos locais. Neste caso nenhum método numérico garante, no estado actual e num tempo de cálculo finito, chegar ao mínimo global da função erro. O mínimo ao qual conduz o algoritmo de aprendizagem é função do seu ponto de partida<sup>165</sup>. Além disso, a acção dos coeficientes do perceptrão sobre as variáveis de saída dependem do estado do perceptrão<sup>166</sup>.

Há dois **modos de utilização dos perceptrões**.

O primeiro é relativo à associação e consiste na reprodução de um conjunto de dados que serviram de aprendizagem ao perceptrão.

O segundo consiste na generalização dos dados desconhecidos apresentados ao perceptrão com a finalidade de emitir um sinal de saída que se espera pertinente. É este modo de utilização que constitui o objecto das nossas preocupações visto levantar o problema da sobre parametragem<sup>167</sup>.

Vários contributos surgiram a este propósito. Weigend e al. (1991) propuseram a cisão dos dados em duas componentes<sup>168</sup> a fim de terminar a aprendizagem em tempo útil. Estes autores propuseram igualmente uma modificação da função erro<sup>169</sup> a fim de penalizar o aumento do número de conexões. Le Curr e al. (1990) propõem a supressão de parâmetros<sup>170</sup> cujo impacto<sup>171</sup> sobre a função

<sup>160</sup> O nível de erro medido pela raiz quadrada do erro quadrado médio.

<sup>161</sup> Devido ao número de parâmetros do modelo.

<sup>162</sup> Contrariamente aos modelos de regressão que supõem uma relação linear e ausência de colinearidade, é suposto somente que os dados são observações de um processo estacionário pelo menos quando se estimarem os parâmetros do modelo.

<sup>163</sup> O número de "níveis" escondidos, o número de células no "nível" escondido, a conexão entre os diferentes "níveis".

<sup>164</sup> O número não está definido.

<sup>165</sup> Inicialização aleatória dos coeficientes do perceptrão.

<sup>166</sup> Nível de outras variáveis.

<sup>167</sup> É suficiente imaginar uma topologia com 5 células de entrada, 10 células sobre o nível escondido (couche cachée), duas células sobre o nível (couche) de saída para que o modelo tenha 70 parâmetros, se o perceptrão é completamente conectado. Nestas condições é necessário dispor de um grande número de dados para proceder à estimação dos parâmetros do modelo, por aprendizagem ou utilizar métodos que reduzam o número de conexões e/ou células do perceptrão.

<sup>168</sup> Uma das componentes é a dos dados que servem de base à aprendizagem e a outra é constituída pelos dados de controle do erro de generalização da aprendizagem.

<sup>169</sup> Os métodos ditos "newtonianos" ou "quase newtonianos".

<sup>170</sup> A supressão de parâmetros é efectuada através do "Optimal Brain Damage".

<sup>171</sup> O impacto é medido pela segunda derivada da função erro.

erro é o menos importante. Cottrel e al. (1996) propõem<sup>172</sup> a generalização da abordagem anterior alargando o teste de student ao contexto não linear.

Outras extensões foram propostas ao binómio "perceptrão - retro-propagação do erro" e dizem respeito à "transfer function" para a qual várias funções foram testadas: as funções radiais, as funções lineares por partes, as funções "step".

Elas incidem sobre a topologia<sup>173</sup> ou sobre a definição de estrutura no seio da rede.

Várias variantes foram propostas ao algoritmo de aprendizagem. As primeiras tentativas dizem respeito aos coeficientes de aprendizagem e à utilização de um coeficiente denominado "moment" tendo por objectivo estabilizar a descida do gradiente.

As avaliações do desempenho<sup>174</sup> (performance) dos diferentes algoritmos conduzem a que os ditos de "segunda ordem" parecem dominar os algoritmos conhecidos por "primeira ordem" como a descida do gradiente.

Há várias tentativas de standardização de algoritmos ditos de "optimização global" tais como o TRUST<sup>175</sup>, cuja performance é considerada boa, mas não garante que o mínimo obtido é o mínimo global ou um dos mínimos globais. Os algoritmos genéticos garantem a obtenção de um mínimo global mas num tempo indeterminado.

#### 4. AS CARTAS AUTO-ORGANIZADAS

O perceptrão "multi-nível" constitui uma categoria de rede de neurónios. Distingue-se as categorias de redes neuronais:

- as recorrentes (feed-back);
- as estruturadas.

As redes recorrentes caracterizam-se pelas suas conexões "para trás", isto é, o valor das células de saída é função, não apenas, do valor das células de entrada e das conexões mas também do estado anterior no qual se encontra a rede.

A utilização da rede, após a aprendizagem, torna-se ela própria um problema de convergência na direcção do estado estável, função do ponto de partida<sup>176</sup>. A utilização prática de este tipo de arquitectura levanta alguns problemas<sup>177</sup> devido ao peso da utilização visto que cada exploração necessita de uma convergência em direcção a um estado estável.

O nosso interesse incide sobre as redes estruturadas e, mais especificamente, sobre uma das suas variantes designada Carta de KOHONEN (1982 e 1995) visto que começa a despertar o interesse em finanças.

A carta de Kohonen<sup>178</sup> pode ser vista como um método de análise de dados que permite ao mesmo tempo reduzir o número de dimensões<sup>179</sup> do problema estudado e reagrupar os indivíduos projectados em perfis tipo<sup>180</sup>, sendo a projecção realizada num espaço discreto.

Uma carta de Kohonen é constituída por um número fixo de unidades, organizadas em uma ou em duas dimensões. Ela contém uma relação física de vizinhança definida<sup>181</sup>. Cada unidade é caracterizada por um vector de parâmetros da mesma dimensão que o espaço de entrada.

O algoritmo é de tipo competitivo e funciona do seguinte modo:

- no momento 0, os parâmetros das unidades são determinados aleatoriamente;

<sup>172</sup> estes autores propõem os "statistical stepwise method for weight elimination".

<sup>173</sup> Através da introdução de conexões "para trás" para formar redes recorrentes cuja fase de utilização levanta o problema da convergência.

<sup>174</sup> Entre estes citamos: Barnard e Holm (1994); Groot e Würtz (1994) e Hgan e Menhaj (1994).

<sup>175</sup> Proposto por Barhen e al. (1993)

<sup>176</sup> O valor das células de entrada.

<sup>177</sup> Apesar do estudo da características da convergência em direcção ao "l'état stable" e da sua convergência.

<sup>178</sup> O perceptrão está para o modelo de regressão tal como a Carta de Kohonen está para a análise em componentes principais.

<sup>179</sup> Isto é, redução do espaço de dados.

<sup>180</sup> Esta classificação não é supervisionada e substitui o método tradicional de classificação tal como a classificação hierárquica.

<sup>181</sup> A cada unidade é afectado um número de ordem que corresponde ao seu lugar físico na carta.

- em cada iteração uma empresa é seleccionada aleatoriamente. Para este indivíduo, a unidade ganhante<sup>182</sup> é determinada. Os parâmetros desta unidade e os seus vizinhos físicos<sup>183</sup> são modificados do modo seguinte:

$$W_i(t+1) = W_i(t) + \Sigma(t)(x(t) - W_i(t)), \text{ para } i \in V_r(i^*)$$

onde:

$W_i(t)$  representa o vector dos parâmetros da unidade  $i$  no instante  $t$ ;

$i^*$  é o índice da unidade ganhante;

$x(t)$  é o indivíduo tirado aleatoriamente no jogo dos dados;

$\Sigma(t)$  é um coeficiente de aprendizagem<sup>184</sup> pequeno e positivo;

$V_r(i)$  é o vizinho físico à volta da unidade  $i$ .

Analisando o processo de aprendizagem constata-se que as unidades são, à partida, afastadas das empresas e progressivamente vão convergindo em direcção ao centro de gravidade de cada grupo.

Na fase de quantificação cada unidade representa o perfil médio de cada grupo.

Passa-se de uma situação caracterizada pela totalidade dos indivíduos da amostra a uma situação onde só há um número reduzido de unidades que representam, cada uma, a média das características dos indivíduos da unidade ganhante.

A segunda fase, a da projecção, é o resultado directo da fase da quantificação e é nela que é efectuado, não somente, o posicionamento das unidades no centro dos grupos, mas também, a organização da sua proximidade no espaço de entrada correspondente ao número de ordem que lhe foi atribuído no espaço de saída<sup>185</sup>.

A ordem física das unidades sobre a grelha respeita a topologia do espaço de entrada.

Quando o parâmetro da vizinhança assume o valor zero<sup>186</sup> o algoritmo de Kohonen transforma-se numa quantificação vectorial<sup>187</sup>.

Quando as entradas terminam o algoritmo de Kohonen converge em direcção ao mínimo local da função de "energia" e generaliza a variância intra classes.

O grau de "finesse" da classificação e a homogeneidade dos grupos é que determina a escolha da estrutura da carta bem como o número de unidades para realizar a quantificação vectorial.

As cartas de Kohonen, relativamente às técnicas clássicas de projecção - a Análise em Componentes Principais- apresentam as seguintes **vantagens**:

1. fornecem uma topologia única do espaço de entrada ao passo que com a ACP pode ser necessário trabalhar com os quatro primeiros eixos que darão lugar a várias representações gráficas;
2. conduzem a um método de classificação em que se pode definir uma noção de vizinhança entre as classes.

As cartas auto-organizadas não deram lugar a tantas publicações como as redes "feed-forward" de tipo perceptrão.

As cartas auto-organizadas são mais uma abordagem de compreensão do que de exploração, isto é, não se trata de explorar a associação linear ou não linear entre variáveis a fim de elaborar modelos de previsão, de aproximação de funções ou de classificação mas de utilizar instrumentos de análise de dados a fim de compreender o que existe numa nuvem de dados multidimensionais. .

<sup>182</sup> A que está mais próxima no sentido da distância euclidiana.

<sup>183</sup> As unidades mais próximas na carta.

<sup>184</sup> É de referir que o coeficiente de aprendizagem e a vizinhança são dois parâmetros do algoritmo que evoluem ao longo da aprendizagem. O coeficiente de aprendizagem decresce progressivamente à medida que as unidades são deslocadas por grandes passos reaproximando-as da sua zona de convergência. A vizinhança evolui igualmente sendo as unidades deslocadas em cada iteração reduzindo-se progressivamente a vizinhança. É de referir que no início dos cálculos,  $\Sigma(t)$  e  $V(t)$  são grandes e, seguidamente, diminuem.

<sup>185</sup> A vizinhança física.

<sup>186</sup> Só a unidade ganhante é deslocada.

<sup>187</sup> Em que a representação é obtida com uma rotação.

As cartas de Kohonen foram recentemente aplicadas no estudo da reestruturação de empresas portuguesas em dificuldades financeiras por Brandão E., de Bodt E. e Levasseur M. (1998). O estudo compreendeu os anos de 1992 a 1996 e uma amostra constituída por 86 empresas dividida em duas sub amostras: sendo 36 empresas em dificuldades e 50 em boa situação financeira. As empresas reestruturadas efectuaram-na segundo as disposições legais portuguesas que regulamentavam esta matéria<sup>188</sup>. Para analisar o impacto do processo de reestruturação os autores serviram-se de 21 variáveis e da análise em componentes principais e das cartas de Kohonen. Os resultados bem sucedidos das empresas reestruturadas não são encorajadores para os legisladores desta matéria.

Estas conclusões foram possíveis graças ao método de análise de dados utilizados - carta de Kohonen - aliada à projecção<sup>189</sup> e à quantificação<sup>190</sup>.

## 5. CONCLUSÃO

Este artigo teve por objecto a apresentação de uma nova metodologia de tratamento de dados em finanças: as redes neuronais. Num primeiro tempo, foram expostos os métodos clássicos, nomeadamente as técnicas estatísticas da regressão múltipla e os métodos dos mínimos quadrados para estimação dos parâmetros. Seguidamente, foram apresentadas as redes neuronais iniciando-se com as recorrentes lineares e não lineares. Por último, foram expostas as cartas auto-organizadas de Kohonen.

As redes de neurónios foram apresentadas como um prolongamento dos métodos estatísticos de tratamento de dados e dois modelos foram retidos: o perceptrão e as cartas de Kohonen.

O perceptrão constitui um prolongamento dos métodos de regressão e foi feita referência à técnica de retro-propagação dos erros (backpropagation) associado ao perceptrão.

As cartas de Kohonen podem ser vistas como uma extensão da análise em componentes principais. Elas oferecem novas possibilidades, combinando as tarefas de projecção e de quantificação, na medida em que permitem compreender melhor o que se passa nas nuvens de dados multidimensionais.

As redes de neurónios<sup>191</sup> não emitem nenhuma hipótese a priori sobre a forma de relação entre as variáveis e permite induzi-las a partir dos dados.

Por último, foi feita referência a um estudo realizado pelo autor, em colaboração, sobre a reestruturação das empresas portuguesas em que foi utilizada a carta de Kohonen.

## BIBLIOGRAFIA

- Altman E.I., Marco G., Varetto F., "Corporate Distress Diagnosis : Comparisons using Linear Discriminant Analysis and Neural Networks", *Journal of Banking and Finance* (18, 1994), p.505-529.
- Brandão E., Levasseur M. e de Bodt E., "Restructuration des Entreprises et Structure Financière", Congresso Internacional da "Connectionist Approches in Economics and Management Sciences", ACSEG, Universidade Católica de "Louvain-la-Neuve", 20/11/1998.
- Coats P.K., Fant L.F., "Recognizing Financial Distress Patterns Using a Neural Network Tool, Financial Management" (vol. 22, Number 3, Autumn 1993), p.142-154.
- Gujarati D.N., *Basic Econometrics*, 3ª ed., New York, McGraw-Hill, 1995.
- Hutchinson J., Lo A.W., Poggio T., "A Nonparametric Approach to Pricing and Hedging Derivative Securities Via Learning Networks", *The Journal of Finance*, (vol. XLIX, nº3, July 1994), p. 851-889.
- Kohonen T., *Self-Organizing Maps*, Springer Series in Information Sciences Vol. 30, Berlin, Springer (1995).

<sup>188</sup> Os decretos leis 177/86, de 2 de Julho e o 132/93, de 23 de Abril.

<sup>189</sup> Que resulta da passagem de um espaço de entrada de 22 dimensões - os rácios utilizados - a um espaço de saída com uma dimensão.

<sup>190</sup> Determinação do perfil tipo dos reagrupamentos das empresas.

<sup>191</sup> As abordagens clássicas emitem hipóteses a priori sobre a forma de relação entre as variáveis.

- Kryzanowski L., Galler M., Wright D.W., "Using Artificial Neural Networks to Pick Stocks", *Financial Analysts Journal* (vol. 49, n°4, July-August 1993), p. 21-27.
- Levasseur M., Quintart A., *Finance*, 3è éd., Paris, Economica, 1998.
- Refenes A.P., *Neural Networks in the Capital Markets*, Wiley, Londres, 1995.
- Refenes A.P., Azema-Barac M., "Neural Network Applications in Financial Asset Management", *Neural Computing and Application* 1994), p. 13-39.
- Ritter H., Martinetz T., Schulten K., *Neural Computation and Self-Organizing Maps, an Introduction*, Addison-Wesley, Reading, 1992.
- Vaga T., *Profiting from Chaos : using choas theory for market timing, stock selection, and option valuation*, New York, McGraw-Hill, 1994.
- Varetto F., Marco G., "Bankruptcy Diagnosis and Neural Networks : Experiments and Comparisons using Linear Discriminant Analysis", Bressanone/Brixen, 1993.
- Weigend A.S., Rumelhart D.E., Huberman B.A., "Generalization by Weight-Elimination", *Advances in Neural Information Processing Systems* (vol. 3, 1991), p. 875-882.

TERCERA PARTE  
CONTABILIDAD 3