

JOSIAS MACIEL

ANÁLISE DE UM SISTEMA DE CRÉDITO COOPERATIVO  
ATRAVÉS DE REDES NEURAS (*MLP*) COM A UTILIZAÇÃO  
DO ALGORITMO LEVENBERG MARQUARDT

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências, no Programa de Pós-Graduação em Métodos Numéricos em Engenharia – Programação Matemática, dos Setores de Tecnologia e Ciências Exatas, UFPR.

Orientadora: Prof<sup>a</sup> Andrea Sell Dyminski, D. Eng

Co-orientadora: Prof<sup>a</sup> Maria Teresinha Arnns Steiner, D. Sc.

CURITIBA  
2005

# TERMO DE APROVAÇÃO

JOSIAS MACIEL

## ANÁLISE DE UM SISTEMA DE CRÉDITO COOPERATIVO ATRAVÉS DE REDES NEURAIS (MLP) COM A UTILIZAÇÃO DO ALGORITMO LEVENBERG MARQUARDT

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre em Ciências, na Área de Concentração em Programação Matemática, do Programa de Pós-Graduação em Métodos Numéricos em Engenharia da Universidade Federal do Paraná, pela comissão formada pelos professores:

Orientadora:

---

Professora Andrea Sell Dyminski, D. Eng.  
Departamento de Pós-Graduação, UFPR.

Co-orientadora:

---

Professora Maria Teresinha Arns Steiner, D. Sc.  
Departamento de Matemática, UFPR.

---

Professor Anselmo Chaves Neto, D. Sc.  
Departamento de Estatística UFPR

---

Professor Christian José Quintana Pinedo, D. Sc.  
Departamento de Matemática, CEFET-PR.

Curitiba, 21 de fevereiro de 2005.

Dedico este trabalho à minha esposa  
GESLAINE e ao meu filho MATHEUS,  
por terem estado ao meu lado de forma  
especial e imprescindível, durante esta  
etapa da minha vida. Pessoas estas que  
me fizeram entender o verdadeiro sentido  
do verbo AMAR...

## **AGRADECIMENTOS**

A Deus, pelo dom da vida, da sabedoria, da perseverança e do amor, dons que alicerçam todas as realizações;

Aos meus pais, João e Shirlei Maciel, e irmãos pelo carinho e incentivo que sempre me transmitiram;

À professora Dr<sup>a</sup> Andréa Sell Dyminski, pela valiosa orientação nas etapas da realização deste trabalho;

À professora Dr<sup>a</sup> Maria Teresinha Arns Steiner, pela disponibilidade em compartilhar seus conhecimentos;

Aos professores Dr. Jair Mendes Marques e Dr. Anselmo Chaves Neto, pela valiosa colaboração prestada com importantes sugestões que contribuíram de forma imprescindível para o desenvolvimento do mesmo;

A Paulino Capelin Fachin, diretor presidente da Cooperativa Agropecuária do Sudoeste do Paraná, pela sua generosa colaboração;

A Eleutério Benin, gerente regional do Sistema de Crédito Cooperativo, que não mediu esforços para fornecer-me informações importantes que possibilitaram o desenvolvimento dos modelos estudados nesta pesquisa;

A Dinarte e Ires Fachin pela confiança e incentivo que sempre me transmitiram;

Ao colega Alex Vieira Falkenberg pelo incentivo e apoio em minha busca pelo saber;

Aos demais professores e colegas do Programa de Mestrado da UFPR - CESEC que, de alguma forma, contribuíram com a realização dessa dissertação.

*Quem espera que a vida seja feita de ilusão pode até ficar maluco ou morrer na solidão. É preciso ter cuidado para mais tarde não sofrer. É preciso saber viver. Toda pedra do caminho você pode retirar. Numa flor que tem espinhos você pode se arranhar. Se o bem e o mal existem você pode escolher. É preciso saber viver.*

*Titãs*

## SUMÁRIO

<b>LISTA DE ABREVIATURAS.....</b>	<b>viii</b>
<b>LISTA DE FIGURAS.....</b>	<b>ix</b>
<b>LISTA DE QUADROS E TABELAS .....</b>	<b>x</b>
<b>CAPÍTULO I.....</b>	<b>1</b>
1.0 – INTRODUÇÃO .....	1
1.1 – CARACTERIZAÇÃO DO PROBLEMA .....	3
1.2 – JUSTIFICATIVA .....	4
1.3 – OBJETIVOS .....	5
1.4 – ESTRUTURA DO TRABALHO .....	6
<b>CAPÍTULO II.....</b>	<b>8</b>
2.0 – REVISÃO DA LITERATURA .....	8
2.1 – ESTRUTURA DE PODER NAS COOPERATIVAS .....	8
2.1.1 – ASSEMBLÉIA GERAL.....	8
2.1.2 – ÓRGÃOS DE ADMINISTRAÇÃO .....	8
2.1.2.1 – CONSELHO DE ADMINISTRAÇÃO.....	8
2.1.2.2 – DIRETORIA EXECUTIVA.....	9
2.1.2.3 – CONSELHO FISCAL.....	9
2.1.3 – MODALIDADES DE COOPERATIVAS DE CRÉDITO .....	9
2.1.3.1 – COOPERATIVA DE CRÉDITO RURAL .....	10
2.1.3.2 – COOPERATIVA DE CRÉDITO MÚTUO .....	10
2.1.4 – OPERAÇÕES ATIVAS, PASSIVAS E ACESSÓRIAS.....	11
2.1.4.1 – OPERAÇÕES ATIVAS .....	11
2.1.4.2 – OPERAÇÕES PASSIVAS .....	12
2.1.4.3 – OPERAÇÕES ACESSÓRIAS .....	12
2.2 – METODOLOGIA PARA DETERMINAÇÃO DO RISCO DE CRÉDITO BASEADA EM TÉCNICAS DE JULGAMENTO.....	14
2.3 – REDES NEURAS ARTIFICIAIS .....	20

2.3.1 – MÉTODO DE OTIMIZAÇÃO .....	21
2.3.1.1 – INTRODUÇÃO .....	21
2.3.1.2 – REDES DE MÚLTIPLAS CAMADAS ALIMENTADAS ADIANTE ( <i>Feed-Foward Multilayer Perceptron</i> ) .....	25
2.3.1.3. – MÉTODO DE LEVENBERG–MARQUARDT .....	32
2.3.2 – SEPARABILIDADE DE PADRÕES .....	35
2.4 – ANÁLISE MULTIVARIADA.....	37
2.4.1 – ANÁLISE DE COMPONENTES PRINCIPAIS .....	37
2.4.2 – ANÁLISE FATORIAL.....	37
2.4.2.1 – MODELO FATORIAL ORTOGONAL.....	38
<b>CAPÍTULO III.....</b>	<b>41</b>
3.0 – MATERIAIS E MÉTODOS.....	41
3.1 – INTRODUÇÃO .....	41
3.2 – BASE DE DADOS .....	42
3.3 – PROCEDIMENTOS DE PREPARAÇÃO DOS DADOS.....	47
3.4 – METODOLOGIA PARA RESOLUÇÃO DO PROBLEMA .....	51
3.4.1 – VALIDAÇÃO CRUZADA.....	53
3.5 – MODELOS DESENVOLVIDOS.....	56
<b>CAPÍTULO IV.....</b>	<b>58</b>
4.0 – OBTENÇÃO DOS RESULTADOS .....	58
4.3 – DESEMPENHO DA <i>MLP</i> .....	60
4.4 – COMPARAÇÃO DE DESEMPENHO DOS MODELOS.....	65
<b>CAPÍTULO V.....</b>	<b>71</b>
5.1 - CONCLUSÕES .....	71
5.2 – SUGESTÕES PARA ESTUDOS FUTUROS.....	74
<b>REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>76</b>
<b>ANEXOS .....</b>	<b>80</b>

## LISTA DE ABREVIATURAS

ACP	Análise de componentes principais
AF	Análise Fatorial
AG	Algoritmo genético
BACEN	Banco Central do Brasil
BP	Retropropagação ( <i>back-propagation</i> )
CMN	Conselho Monetário Nacional
LMS	Mínimo quadrado médio ( <i>least-mean-square</i> )
LM	Levenberg-Marquardt
<i>MLP</i>	Perceptron de múltiplas camadas ( <i>multilayer perceptron</i> )
MSE	Erro médio quadrático ( <i>Mean squared error</i> )
<i>MLF</i>	<i>Multilayer feedforward</i>
RN	Rede neural
RMSE	Raiz quadrada do erro médio quadrático
SICREDI	Sistema Integrado de Crédito Cooperativo
SPC	Serviço de Proteção ao Crédito
VC	Validação cruzada
V-C	Vapnik-Chervonenkis (dimensão de)
VCG	Validação cruzada generalizada

## LISTA DE FIGURAS

FIGURA 2.1 - Localização geográfica da região de estudo .....	13
FIGURA 2.2 - Os C's do Crédito .....	15
FIGURA 2.3 – Exemplo de Rede Neural Multicamadas <i>Feedforward</i> , totalmente conectada.....	26
FIGURA 2.4 – Modelo não-linear de um neurônio .....	27
FIGURA 2.5 – (a) Função de limiar; (b) Função linear por partes; (c) Função sigmóide para parâmetro de inclinação $a$ variável.....	28
FIGURA 2.6 – Ilustração da aprendizagem por correção de erro para o neurônio $k$ . .....	32
FIGURA 3.1 – Fatores X Autovalores da análise realizada .....	49
FIGURA 3.2 – Treinamento versus capacidade de generalização da RN.....	55
FIGURA 4.1 – Ilustração para a identificação do melhor valor de corte, neste caso para a rede 9E16n88p.....	61
FIGURA 4.2 – Ilustração da performance da rede na fase de treinamento. Neste caso, da rede 9Et8n70t. ....	62
FIGURA 4.3 – Ilustração da performance da rede na fase de teste. Neste caso, da rede 9Et8n70t.....	63
FIGURA 4.4 – Ilustração da performance da rede na fase de treinamento. Neste caso, da rede 9Et10n65l. ....	65
FIGURA 4.5 – Ilustração da performance da rede na fase de teste. Neste caso, da rede 9Et10n65l.....	65
FIGURA 4.6 – Acertos das classes de saída por arquitetura. ....	67
FIGURA 4.7 – Ilustração da performance no desempenho da rede. Neste caso, da rede 9Et10n65l.....	68
FIGURA 4.8 – Ilustração da performance no desempenho da rede. Neste caso, da rede 9Et8n70t.....	69

## LISTA DE QUADROS E TABELAS

QUADRO 2.1 - Balanço Hipotético.....	17
QUADRO 2.2 – Fontes de Receitas e Despesas .....	17
QUADRO 2.3 – Algoritmo de Levenberg–Marquardt.....	34
QUADRO 3.1 – Principais distinções entre Banco e Cooperativa de Crédito. ....	41
QUADRO 3.2 – Classificação de Risco.....	43
QUADRO 3.3 – Classificação quanto ao atraso no pagamento.....	43
QUADRO 3.4 – Campos selecionados e valores adotados. ....	45
QUADRO 3.5 – Percentuais referentes às observações.....	47
TABELA 3.1 – Autovalores em ordem decrescente de magnitude.....	48
TABELA 3.2 – Matriz Fatorial (antes da rotação).....	50
TABELA 4.1 Percentuais de erros obtidos nas simulações para um n° de variável de neurônios ocultos .....	59
TABELA 4.1 Percentuais de erros obtidos nas simulações para um n° de variável de neurônios ocultos .....	60

## RESUMO

Os problemas envolvendo tomada de decisões estão presentes em todas as áreas da atividade humana e, ao longo da história, foram tratados através de uma série de abordagens diferentes. Este estudo propõe uma abordagem estruturada na exploração das Redes Neurais Artificiais (RNA's) como instrumento de apoio à previsão de inadimplência de pessoa física, utilizando o banco de dados de uma Cooperativa de Crédito da região do sudoeste do Paraná, que contém informações de seus associados/clientes. Face a necessidade de se tomar decisões bem estruturadas e com base no histórico de pessoas físicas, o modelo de avaliação do risco de inadimplência foi construído usando dados disponíveis da cooperativa de crédito, a qual possui tanto informações sobre os associados adimplentes, quanto dos inadimplentes. A metodologia do estudo propõe inicialmente, uma avaliação estatística dos dados tanto pela análise de componentes principais quanto pela análise fatorial, que teve o propósito de avaliar a importância relativa das diferentes variáveis envolvidas no processo de análise de crédito. As redes neurais desenvolvidas foram do tipo feed-forward, totalmente conectadas. Para se determinar os “melhores” pesos de treinamento foi utilizado o algoritmo de Levenberg Marquardt. Finalmente, o sucesso da modelagem foi avaliado através de uma amostra de controle, utilizada para testar a capacidade de classificação das redes neurais, sendo esta amostra diferente da usada para obter o modelo que decidirá a respeito da concessão (ou não) do crédito ao associado/cliente.

## **ABSTRACT**

The problems that evolve how to get decisions are present in all the areas and, along the history, they were treated through a series of different discussions. This application suggest an discussion that have structure in the exploration of the “artificial neural network (ANN)” like an instrument to give support to the estimate of the risks caused by the fysical person adimplence, using the information bank of a Credit Co-operative in the South West Paraná region. That has information about it associates/clients. The obtainment of the information, that compose an offer of a credit concession, has to be versatile and assorted, composite by information that are important in the process of credit analysis, what means that it consent evaluate the grade of the risk that it have in the moment that it gives credit to the client. In face of the necessity to take decisions with good structures and based in the description of the fysical person, the model of the adimplence risk estimate will be constructed using informations desponible by the credit co-operative, that have information about the adimplentes and inadimplentes associates. It will be selected variable through the importance of wich one in the inadimplence risk estimate to the training of the net. Finally, the success of the modality will be appraised throught an control sample, that will be used to test the capacity of classification of the neural network, beinz this sample different of the another one used to get the model that will decide the concession or not of the credit to the associate/client.

## CAPÍTULO I

### 1.0 – INTRODUÇÃO

A palavra “crédito” pode ter mais de um significado, dependendo do contexto sob o qual esteja sendo tratada [GUIMARÃES, 2000]. Do ponto de vista meramente empresarial, a concessão de crédito significa a transferência da posse de um bem, ou de uma quantia em dinheiro, mediante a promessa de devolução no futuro. De acordo com esse conceito, pode-se entender o crédito à disposição de uma pessoa, física ou jurídica, como a capacidade da mesma em obter dinheiro, produtos ou serviços mediante compromisso de pagamento num determinado período de tempo.

Em toda e qualquer atividade que uma instituição financeira venha a atuar, sempre haverá um componente maior ou menor de risco [STEINER *et al.*, 2003]. Esse risco nasce da probabilidade que um evento desejado venha a não ocorrer. Desta forma, o risco como enfoque estatístico se fundamenta em dados históricos, permitindo que as decisões sejam tomadas a partir de estimativas consideradas aceitáveis. Assim, quando é concedida uma operação de crédito, espera-se receber, em uma data futura, não só o principal como também seus rendimentos. Assim, crédito, como a própria origem da palavra expressa (*crederer = acreditar*), é a crença de que o cliente honrará os compromissos para com seu credor. Há fatores que contribuem para elevar o risco das operações de crédito. Tais fatores, se não cuidadosamente administrados, podem reduzir significativamente a qualidade do crédito [ARRAES *et al.*, 2001].

A política de crédito, de uma instituição, é o principal meio através do qual são direcionadas as atividades de crédito, buscando-se o maior grau de eficiência na gestão de tais riscos. Além de estabelecer padrões, a política de crédito estabelece, também, a filosofia básica de crédito desta instituição. A partir dela, são traçados e buscados os objetivos de qualidade das carteiras<sup>1</sup> e dos resultados, delimitando os

---

1 – Carteira é o conjunto de contratos de uma mesma operação com as mesmas características de tratamento.

níveis de tolerância ao risco, orientando as atividades de crédito de uma maneira coerente com a estratégia da instituição, pois uma instituição financeira, como instrumento de organização econômica, tem a responsabilidade de preservar os recursos que a ela são confiados. A adequada gestão dos recursos deve gerar condições para o atendimento das demandas de seus clientes por produtos de crédito.

As Cooperativas de Crédito são componentes do Sistema Financeiro Nacional e seguem orientação das políticas estabelecidas pelo Conselho Monetário Nacional e pelo Banco Central [BANSICREDI, 2002]<sup>2</sup>. Elas captam recursos de seus clientes e do mercado financeiro, repassando-os, posteriormente, sob a forma de empréstimos e/ou financiamentos, na expectativa de realizarem pequenos ganhos aos seus associados. Neste processo, o risco total corresponde à condição de não-recebimento dos empréstimos, face à concessão de financiamentos para tomadores de alto risco. GURLEY e SHAW, *apud* SANTOS 2000, destacam, como importante regra a ser seguida pelas instituições financeiras na atividade de concessão de empréstimos, a atenção que deve ser direcionada para a análise de compatibilidade da “capacidade financeira” dos tomadores com o cronograma de amortização do empréstimo.

Em meio às considerações acima explicitadas sobre a realidade presente nas cooperativas de crédito, propõe-se uma análise junto às instâncias envolvidas no Sistema Financeiro, quer estejam voltadas à agência financiadora, quer aos clientes da mesma. O enfoque deste estudo centra-se na adimplência (ou não) de pessoa física, a qual será abordado a partir da exploração de Redes Neurais Artificiais como forma de minimizar os fatores de riscos presentes neste tipo de relação financeira.

---

2 – BANSICREDI, **Manual de Políticas de Crédito** – Edição: Bansicredi – Gerência de Crédito. Abril (2002).

## 1.1 – CARACTERIZAÇÃO DO PROBLEMA

O gerenciamento do risco de fornecer (ou não) o crédito financeiro aos associados do Sistema Integrado de Crédito Cooperativo, bem como qual o montante a ser disponibilizado, é hoje uma necessidade.

O desenvolvimento ou aprimoramento de novas tecnologias, baseadas em "Sistemas de apoio à decisão", que tem como intuito a busca pelo aperfeiçoamento da análise de viabilidade para concessão de crédito, tanto para o associado/cliente quanto para a cooperativa, são os alicerces para que este gerenciamento se concretize de forma satisfatória para ambos [SICREDI, 2003]<sup>3</sup>.

Os problemas envolvendo tomada de decisões estão presentes em todas as áreas e, ao longo da história, foram tratados através de uma série de abordagens diferentes [BUSCEMA e SACCO, 2003; ATIYA, 2001; YANG *et al*, 2001; WONG e SELVI, 1998]. Mesmo antes de se falar em “sistemas de apoio à decisão” e antes ainda dos computadores, as mesmas motivações, que hoje impulsionam áreas como a chamada inteligência do negócio e o surgimento de novas tecnologias, já existiam. As decisões para concessão de crédito a clientes são anteriores aos computadores e diversas gerações de sistemas foram implementadas para solucioná-los, com o intuito de auxiliar os gestores de crédito na tomada de decisão. As Redes Neurais Artificiais, como modelo decisório, têm sido utilizadas em um grande número de áreas de negócios, tais como: Administradoras de cartões de crédito, Administradoras de Planos de Saúde, Bancos, Financeiras, Seguradoras e outros, que envolvem desde investimentos de grandes montantes, ou seja, aplicações em bolsas de valores, até empréstimos de menores vultos, como o financiamento a Pessoa Jurídica [STEINER *et al*, 2003; BACK e CHO, 2003; FRANCO, 2001; ARRAES *et al*, 2001; SANTOS, 2000; CARDOSO, 2000; STEINER *et al*, 1999].

---

3 – SICREDI-2003, **Legislação e Formalística de Crédito**, Sistema Integrado de Crédito Cooperativo.

Diversas tecnologias se sucederam no apoio aos processos decisórios. A necessidade de se tomar decisões, bem estruturadas, motivou a utilização da estatística e a criação de ferramentas para sumarizar informações e apresentá-las a gerentes e executivos, provendo excelentes oportunidades de aplicação de tecnologias derivadas da pesquisa em processamento paralelo.

Desta forma, tem-se como problema a ser resolvido, o seguinte questionamento: “Como obter um sistema adequado de tomada de decisão, a fim de se evitar perdas para o Sistema de Crédito Cooperativo, tanto pela concessão de crédito a maus cooperados, como pela rejeição de crédito a bons cooperados?”

## **1.2 – JUSTIFICATIVA**

O Brasil tem um número elevado de clientes inadimplentes<sup>4</sup> no sistema de crédito financeiro, segundo levantamento da SERASA (maio, 2003), que se destaca por ser uma das maiores empresas do Brasil em informações e análises econômico-financeiras para apoiar decisões de crédito e negócios. Por ocasião da coleta de dados no início de 2004, seu indicador “Serasa de Inadimplência”, ressalta que a inadimplência total (pessoa física e jurídica) apresentou crescimento de 3,2% no primeiro quadrimestre de 2003, na comparação com o mesmo período de 2002. O indicador também apontou queda do ritmo de crescimento de pessoa física. De janeiro a abril de 2003, a inadimplência de pessoa física cresceu 7% na comparação com o mesmo período de 2002.

Segundo o presidente da SERASA, Sr. Elcio Anibal de Lucca, no caso de pessoa física, a falta de metodologia adequada para concessão de crédito juntamente com o agravamento da conjuntura para o consumidor – juros elevados, acentuação da queda da renda e aumento do desemprego - foram determinantes para a inadimplência na concessão de crédito financeiro.

---

4 – Inadimplência significa que um contrato está em atraso superior ao prazo máximo estabelecido legalmente para pagamento.

Por essas razões, a elevação sazonal da inadimplência que normalmente se limitava ao 1º trimestre, em 2004 se estendeu ao longo do 1º semestre, mesmo com fraca demanda por crédito, com o consumidor retraído e com baixa atividade econômica.

As Redes Neurais Artificiais têm sido mencionadas por alguns autores [BUSCEMA e SACCO, 2003; FRANCO, 2001; ATIYA, 2001; STEINER *et al.*, 1999 e outros] como uma ferramenta bastante poderosa na administração do risco de inadimplência existente na concessão de crédito. Especificamente, cabe citar o trabalho de FRANCO (2001), que faz uso de um sistema neural para avaliação de crédito a pessoa física no acompanhamento da carteira. O gerenciamento do risco de fornecer (ou não) o crédito financeiro ao associado de uma cooperativa de crédito, bem como o montante a ser disponibilizado é, hoje, uma necessidade. O desenvolvimento ou aprimoramento de novas tecnologias baseadas em "Sistemas de apoio à decisão", que tem como intuito a busca pelo aperfeiçoamento da análise de viabilidade para concessão de crédito, tanto para o associado quanto para a cooperativa, é o alicerce para que este gerenciamento se concretize de forma satisfatória para ambos.

### **1.3 – OBJETIVOS**

Este estudo propõe uma abordagem estruturada para exploração das Redes Neurais Artificiais como instrumento de apoio à avaliação de riscos de inadimplência, que permitirá sistematizar o processo de gestão de empréstimos do Sistema Integrado de Crédito Cooperativo de uma forma totalmente integrada e baseado numa filosofia de otimização de processos, avaliando o setor de uma cooperativa de crédito na análise de concessão de crédito a seus associados. Especificamente, neste caso, as redes “aprendem” a identificar o risco de um certo associado (pessoa física) ser adimplente (ou não), ou seja, honrar com o pagamento do empréstimo e/ou financiamento adquirido.

O estudo desenvolvido segue o estabelecido na Resolução nº 2.682, 21 de dezembro de 1999, do Conselho Monetário Nacional e com base no Manual de Políticas de Crédito, do Sistema Integrado de Crédito Cooperativo, que apresenta as diretrizes básicas relativas aos aspectos relacionados à gestão dos créditos existentes nas carteiras de empréstimos, o qual destaca que o grau de risco de cada operação é apurado de acordo com a classificação de risco do tomador e suas características e da avaliação do grau de risco da operação.

Em face desta necessidade de gerenciamento do risco, o objetivo deste trabalho é a pesquisa para obtenção de modelos de Redes Neurais Artificiais de Múltiplas Camadas, capazes de classificar clientes do Sistema Integrado de Crédito Cooperativo, determinando uma fronteira entre clientes adimplentes e inadimplentes, uma vez que o sucesso do empréstimo, e/ou financiamento, é essencial para a existência da cooperativa de crédito, o qual gerará a captação de novos negócios e, conseqüentemente, novos associados/clientes.

#### **1.4 – ESTRUTURA DO TRABALHO**

O presente trabalho é composto por cinco capítulos: Introdução, Revisão da Literatura, Materiais e Métodos, Resultados e Conclusão.

No capítulo reservado à revisão de literatura são apresentadas as técnicas teóricas usadas nesse trabalho de análise do risco de inadimplência em carteiras de empréstimos de pessoas físicas. Primeiramente apresenta-se a metodologia para determinação do risco de crédito para pessoa física. A seguir abordam-se as redes neurais artificiais, com uma breve revisão sobre neurônios artificiais, sendo dada ênfase à rede de múltiplas camadas (*multilayer perceptron - MLP*) treinada de forma supervisionada através do tradicional algoritmo de retropropagação de erro (*error back-propagation*). Posteriormente apresenta-se o algoritmo de Levenberg–

Marquardt, que é considerado uma otimização do algoritmo *back-propagation*. A revisão é concluída com a apresentação de duas técnicas de Análise Multivariada: Análise de Componentes Principais e Análise Fatorial.

No terceiro capítulo, Materiais e Métodos, são apresentadas as variáveis usadas na formação do banco de dados, os valores assumidos pelas mesmas e o resultado da análise de componentes principais e análise fatorial. Além disso, são comentadas as características destas variáveis, bem como o método utilizado para avaliar a eficiência dos procedimentos.

O quarto e o quinto capítulo apresentam, respectivamente, os Resultados e as Conclusões. Na parte dos resultados, são comparadas as taxas de acerto obtidas pelas redes construídas com as do modelo usado atualmente pela cooperativa de crédito. Também é levantada a necessidade de reformulação do questionário para coleta de informações, de modo que se possa obter um conjunto de dados não apenas confiável, mas que também ofereça condições para a obtenção de modelos eficientes, no sentido discriminatório.

## **CAPÍTULO II**

### **2.0 – REVISÃO DA LITERATURA**

### **2.1 – ESTRUTURA DE PODER NAS COOPERATIVAS**

As cooperativas têm uma estrutura de poder disciplinada pela Lei nº 5.764/71.

#### **2.1.1 – ASSEMBLÉIA GERAL**

*Assembléia Geral, nos termos do artigo da Lei 5.764/71, “a Assembléia Geral dos associados é o órgão supremo da sociedade, dentro dos limites legais e estatutários, com poderes para decidir os negócios relativos ao objeto da sociedade e tomar as resoluções convenientes a desenvolvimento e defesa desta, e suas deliberações vinculam a todos, ainda que ausentes ou discordantes”.*

As Assembléias podem ser de:

- Constituição;
- Ordinária;
- Extraordinária;
- Liquidação.

#### **2.1.2 – ÓRGÃOS DE ADMINISTRAÇÃO**

##### **2.1.2.1 – CONSELHO DE ADMINISTRAÇÃO**

É órgão colegiado composto exclusivamente de associados eleitos em Assembléia Geral para administrar a sociedade, traçando as metas e os objetivos a serem cumpridos. A posse dos eleitos depende da homologação de seus nomes pelo BACEN – artigo 1º, parágrafo segundo da Resolução do Conselho Monetário

Nacional nº 3.014, de 28/12/2002. No Paraná, o Conselho é composto por: Presidente, Vice-Presidente e no mínimo 3 e no máximo 10 Conselheiros, conforme previsto no Estatuto Social.

#### **2.1.2.2 – DIRETORIA EXECUTIVA**

É o órgão que visa executar, diretamente, os atos de gestão da Cooperativa, composto pelo Presidente e pelo Vice-Presidente.

#### **2.1.2.3 – CONSELHO FISCAL**

É um órgão autônomo, desvinculado de qualquer subordinação dos órgãos de administração e que tem a atribuição de fiscalizar a sociedade. O Conselho Fiscal é composto por 6 (seis) membros, todos associados da cooperativa, sendo três membros efetivos e três membros suplentes e são eleitos em Assembléia Geral, tomando posse somente após terem os seus nomes homologados pelo Banco Central do Brasil.

#### **2.1.3 – MODALIDADES DE COOPERATIVAS DE CRÉDITO**

A admissão no quadro social de cooperativa de crédito está estreitamente vinculada a modalidade de cooperativa constituída.

Assim, consoante a Lei nº 5.764, de 16/12/1971 e Regulamento anexo à Resolução CMN nº 2.771, de 30/08/2000, com as alterações da Resolução CMN nº 3.058, de 20/12/2002, podem ser constituídas cooperativas de crédito rural e cooperativas de crédito mútuo, observando-se sempre as regras abaixo mencionadas:

### **2.1.3.1 – COOPERATIVA DE CRÉDITO RURAL**

Constituída por pessoas físicas que desenvolvam, na área de atuação da cooperativa, de forma efetiva e predominante, atividades agrícolas, pecuárias ou extrativas, ou se dediquem a operações de captura e transformação do pescado.

Podem admitir no quadro social:

- a) Empregados da própria cooperativa de crédito, das entidades a ela associadas e daquelas de cujo capital participem, e pessoas físicas prestadoras de serviços, em caráter não eventual, a cooperativa de crédito e as referidas entidades, equiparados aos primeiros no tocante aos seus direitos e deveres como associados. (não podem votar nem serem votados – Art. 31 da Lei 5.764/71);
- b) Aposentados que, quando em atividade, atendiam aos critérios estatutários de associação;
- c) Pais, cônjuge ou companheiro, viúvo e dependente leal de associado e pensionista de associado falecido;
- d) Pessoas jurídicas que exerçam exclusivamente atividades agrícolas, pecuárias ou extrativas na área de ação da cooperativa ou atividades de captura ou transformação do pescado;
- e) Pessoas jurídicas sem fins lucrativos.

### **2.1.3.2 – COOPERATIVA DE CRÉDITO MÚTUO**

Pode ser constituída por:

- a) Empregados ou servidores, e prestadores de serviço em caráter não eventual de:
  - 1. determinada entidade pública ou privada;
  - 2. determinado conglomerado econômico;
  - 3. conjunto definido de órgãos públicos, hierárquica ou administrativamente vinculados;

4. conjunto definido de pessoas jurídicas que desenvolvam atividades idênticas ou estreitamente correlacionadas por afinidade ou complementaridade;

b) Trabalhadores de:

1. determinada profissão regulamentada;
2. determinada atividade, definida quanto à especialização;
3. conjunto definido de profissões ou atividades cujos objetos sejam idênticos ou estreitamente correlacionados por afinidade (mesmos interesses) ou complementaridade;

c) *pequeno empresário, microempresário ou microempreendedor*, responsável por negócio industrial, comercial ou de prestação de serviços, inclusive as atividades da área rural (pessoas que desenvolvam, na área de atuação da cooperativa, de forma efetiva e predominante, atividades agrícolas, pecuárias ou extrativas, ou se dediquem a operações de captura e transformação do pescado), seja igual ou inferior ao limite estabelecido pela legislação em vigor para as pequenas empresas. Para que este tipo de cooperativa seja autorizada a funcionar pelo Banco Central do Brasil, deverá cumprir exigências a mais estabelecidas pela Resolução CMN nº 3.058, 2002.

#### **2.1.4 – OPERAÇÕES ATIVAS, PASSIVAS E ACESSÓRIAS**

De conformidade com o Regulamento anexo à Resolução CMN nº 2.711, de 30/08/2000, as cooperativas de crédito podem praticar as seguintes operações:

##### **2.1.4.1 – OPERAÇÕES ATIVAS**

Operações ativas são aquelas em que a cooperativa se torna credora do associado, em virtude de concessão de créditos, nas modalidades de:

- a) desconto de títulos;

- b) operações de empréstimos e de financiamento;
- c) crédito rural;
- d) repasse de recursos oriundos de órgãos oficiais e de instituições financeiras.

#### **2.1.4.2 – OPERAÇÕES PASSIVAS**

Operações passivas são aquelas que importarem em ônus e obrigações para a cooperativa, que, na relação jurídica, se torna devedora do associado, nas seguintes modalidades:

- a) captação de depósitos à vista;
- b) captação de depósitos a prazo, sem emissão de certificado;
- c) captação de recursos de instituições financeiras, nacionais e estrangeiras, na forma de empréstimos, repasses, refinanciamentos e outras modalidades de operações de crédito;
- d) captação de recursos de qualquer entidade, na forma de doações, de empréstimos ou repasses, em caráter eventual, isentos de remuneração ou a taxas favorecidas.

#### **2.1.4.3 – OPERAÇÕES ACESSÓRIAS**

Operações acessórias são aquelas que não importam na concessão de crédito mas em serviços prestados pela cooperativa, tais como:

- a) de cobrança, de custódia, de correspondente no País, de recebimento e pagamentos por conta de terceiros e sob convênio com instituições públicas e privadas, nos termos da regulamentação aplicável às demais instituições financeiras;

- b) a outras instituições financeiras, mediante convênio, para recebimento e pagamento de recursos coletados com vistas a aplicação em depósitos, fundos e outras operações disponibilizadas pela instituição conveniente;
- c) formalização de convênios com outras instituições financeiras com vistas a:
  - i) obter acesso indireto a conta Reservas Bancárias, na forma da regulamentação em vigor;
  - ii) participar do Serviço de Compensação de Cheques e Outros Papéis (SCCP);
  - iii) realizar outros serviços complementares as atividades fins da Cooperativa.

A localização geográfica da região de estudo, sudoeste do Paraná, é apresentada através da Figura 2.1 que dista aproximadamente 450km da capital.

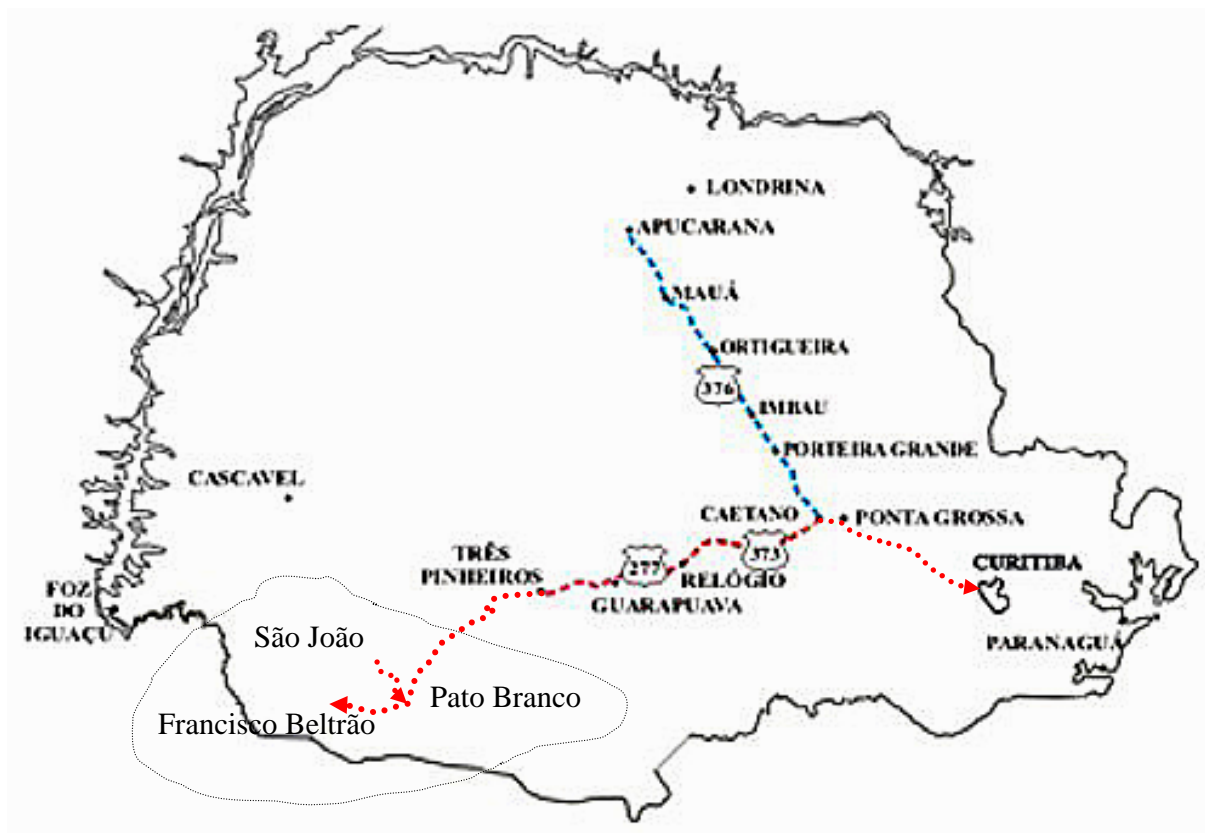


FIGURA 2.1 - Localização geográfica da região de estudo

## **2.2 – METODOLOGIA PARA DETERMINAÇÃO DO RISCO DE CRÉDITO BASEADA EM TÉCNICAS DE JULGAMENTO**

GITMAN (1997) destaca que um dos fatores básicos para decisão final de crédito é o julgamento subjetivo que deve ser realizado a fim de determinar se é válido ou não assumir riscos. Por definição, o processo de análise subjetivo envolve decisões individuais para a concessão ou recusa de empréstimos. Neste processo, a decisão se baseia na experiência adquirida, disponibilidade de informações e sensibilidade de cada analista quanto à viabilidade do crédito. Ainda, segundo o autor, a experiência adquirida do analista e a disponibilidade de informações (internas e externas) sobre o caráter do tomador são requisitos fundamentais para a análise do risco de crédito.

Para STEINER *et al*, (1999), qualquer erro na decisão de concessão de crédito pode significar que em uma única operação haja a perda do ganho obtido em dezenas de outras bem-sucedidas. Assim, é importante prever a inadimplência, pois os prejuízos com operações mal sucedidas provavelmente serão cobertos com a cobrança de uma sobretaxa a novos créditos, ou seja, encarece-se as operações futuras.

Na análise de risco são levados em consideração alguns critérios, sendo que os fatores de risco têm-se baseado em informações que são ditas como necessárias para a análise subjetiva da capacidade financeira dos tomadores, que são tradicionalmente conhecidas como os “C’s do Crédito” [CARDOSO, 2000; LIMA 2002]. Tendo como objetivo apresentar os C’s de forma resumida, a Figura 2.2 mostra os mesmos com uma breve descrição.

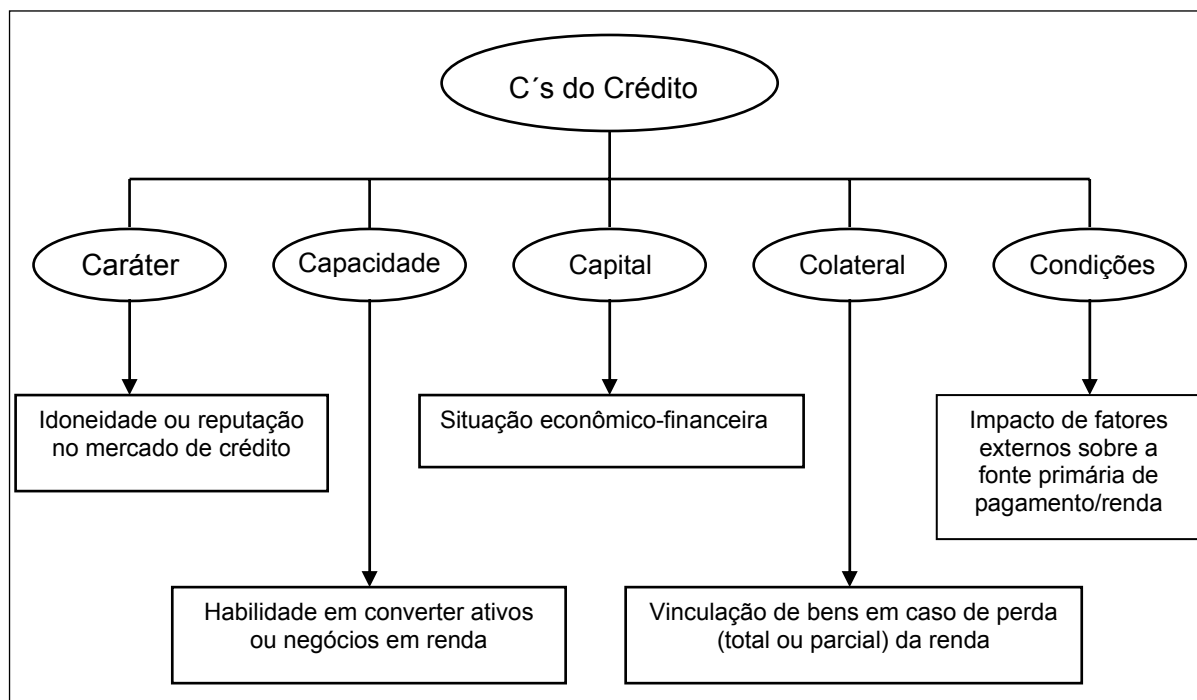


FIGURA 2.2 - Os C's do Crédito

FONTE: SANTOS, 2000.

O critério CARÁTER está relacionado com a probabilidade de que os tomadores de crédito amortizem seus empréstimos. Nesta análise é indispensável que existam informações históricas dos tomadores quanto à intencionalidade e pontualidade na amortização de empréstimos. Mesmo considerando esses aspectos não fica descartada a possibilidade de a qualquer momento o credor deixar de ter esse “caráter”, ou seja, mudar a sua forma de agir em relação aos credores ao longo do tempo. Esta preocupação é apresentada por SOUZA *et. al* (2000), quando cita o ditado popular: “alguém é honesto até o dia em que o deixa de ser”.

Internamente, as informações do caráter dos tomadores podem ser extraídas do fluxo de movimentação da conta corrente, composto pelo histórico de créditos e débitos. Para CHAKRAVARTY e SCOTT, *apud* SANTOS (2000), quanto maior for o tempo ou a durabilidade da conta corrente, maior será a possibilidade em adquirir conhecimento sobre o caráter de seus clientes. Externamente, o acesso ao banco de dados de empresas especializadas no gerenciamento de informações de

crédito (ex: Serasa S.A., Equifax do Brasil Ltda) podem proporcionar informações evidenciadoras do caráter dos tomadores.

O critério CAPACIDADE refere-se ao julgamento subjetivo do analista quanto à habilidade dos tomadores de empréstimos no gerenciamento e conversão de seus negócios em renda. Identifica-se a renda como a “fonte primária de pagamento” e, portanto, como o principal referencial para possibilitar a análise da compatibilidade do empréstimo com a capacidade financeira do tomador [CARDOSO, 2000]. A identificação da renda total do tomador e posterior análise de compatibilidade com o empréstimo solicitado é uma tarefa indispensável para a tomada de decisão. BLACK e MORGAN, *apud* Santos (2000), argumentam que deve ser dada atenção especial à análise da renda, uma vez que consideram existir uma relação direta entre a perda de renda e a taxa de inadimplência de pessoas físicas em empréstimos. LIMA (2002), define a capacidade como sendo o potencial do cliente, salientando a importância deste fator e destacando como abrangência: competências, competitividade e aspectos técnicos do negócio a ser financiado.

Alguns autores têm dado uma conotação ao fator capacidade no sentido de capacidade de pagamento. Para SOUZA *et. al.* (2000), a capacidade deve estar relacionada aos fatores que contribuem para a empresa ser competente e competitiva. Isto facilitará sua condição de ser capaz de pagar suas dívidas, mas não é a condição de pagamento. LIMA (2002) destaca como exemplo. “É como um meio de transporte que nos leva a um lugar, mas não é o lugar”. Daí, a capacidade de pagamento deve ser examinada por meio de análise financeira.

A riqueza do tomador é representada pelo patrimônio líquido e é determinada através da diferença entre o ativo total (recursos/renda) e o passivo total (dívidas). Verifica-se que caso o valor de mercado dos ativos seja maior que o valor de mercado dos passivos, há um aumento confirmado da riqueza do tomador, face à maior disposição de recursos (financeiros e patrimoniais) geradores de renda.

O Quadro 2.1 apresenta uma estrutura hipotética com possíveis contas, componentes de um balanço patrimonial de pessoa física.

ATIVOS	PASSIVO E PATRIMÔNIO LÍQUIDO
<b>Ativos Circulantes</b> Dinheiro disponível e/ou em conta corrente Dinheiro em aplicações financeiras Salário a Receber  <b>Ativos Conversíveis em Prazos Superiores a 12 Meses</b>  <b>Ativos em Investimentos:</b> Bens alimentícios Vestuários Móveis e utensílios Veículos Ações de empresas Outros Ativos  <b>Total do Ativo</b>	<b>Passivos Circulantes</b> Crediário Empréstimos bancários Mensalidade escolar Planos de saúde Aluguel residencial Impostos e taxas Outros  <b>Passivos Amortizáveis em Prazos Superiores a 12 Meses</b> Financiamentos imobiliários Financiamentos de veículos Outros  <b>Patrimônio Líquido</b>  <b>Total do Passivo e Patrimônio Líquido</b>

QUADRO 2.1 - Balanço Hipotético

FONTE: SANTOS, 2000.

BARRON e STATEN, *apud* Santos (2000), propõem as seguintes fontes geradoras de receitas e despesas às pessoas físicas:

FONTES GERADORAS DE RECEITAS	FONTES GERADORAS DE DESPESAS
<b>Receitas com:</b> Atividade assalariada Aluguel de imóveis Aplicações financeiras Pensões Heranças Seguros Atividade autônoma, liberal e/ou empresarial	<b>Despesas com:</b> Alimentação Aluguel da moradia Educação Saúde Lazer Aquisição de bens e serviços Empréstimos bancários
<b>Saldo Líquido Mensal = Total das Receitas – Total das Despesas</b>	

QUADRO 2.2 – Fontes de Receitas e Despesas

FONTE: SANTOS, 2000.

O critério CAPITAL é determinado pela situação econômica, patrimonial e financeira do tomador, observando-se a composição dos recursos (qualitativa e quantitativa), onde são aplicados e como são financiados [CARDOSO, 2000]. Esta informação de crédito da pessoa física está diretamente relacionada com a capacidade de geração de renda para cobrir os custos do empréstimo. Por isso, para avaliá-la, são utilizados os procedimentos para a análise de capacidade.

O critério COLATERAL está associado com a análise da riqueza patrimonial de pessoas físicas (bens móveis e imóveis), considerando a possibilidade futura de vinculá-la (total ou parcialmente) ao contrato de empréstimo, em caso de perda da fonte primária de pagamento (renda). BARRON, *apud* Santos (2000), destaca a importância inibidora da vinculação de bens patrimoniais em contratos de empréstimos, ao considerar que esta condição torna os tomadores mais propensos à amortização pontual do empréstimo, para que possam reaver os bens exigidos pela concessão do crédito como garantia. Para SCHRICKEL (1998) o colateral configura na garantia, tendo o papel de contrabalançar o resultado negativo em um dos outros elementos. Segundo os autores WESTON e BRIGHAM, *apud* LIMA (2002), o colateral é representado pelos ativos que o cliente pode oferecer como garantia para o crédito estendido a ele.

No critério CONDIÇÕES, a análise de crédito baseia-se em informações relacionadas com as condições externas e está associada com o impacto de fatores sistemáticos sobre a fonte primária de pagamento (renda) dos tomadores. A atenção nesta informação é de extrema importância para a determinação do risco total de crédito, uma vez que, dependendo da gravidade do fator sistemático (ex: aumento da taxa de desemprego), a cooperativa de crédito poderá enfrentar grandes dificuldades para receber os empréstimos.

No processo decisório do crédito, geralmente, o agente financiador tem como parâmetros alguns fatores ao definir sobre a concessão ou não do crédito e o seu montante. Para SCHRICKEL (1998) pode-se apresentar uma subdivisão entre os C's do crédito:

- Aspectos pessoais: caráter e capacidade;
- Aspectos financeiros: capital e condições.

Para o autor o “C” colateral deve ser considerado, ou exigido, quando os C’s do aspecto financeiro, capital e condições, não forem suficientes para a decisão do crédito.

Destaca-se como primeira etapa da análise de crédito, a pesquisa do “caráter” do cooperado/cliente. Caso existam informações desabonadoras do tomador com o mercado de crédito, a decisão tende a ser a imediata recusa ao empréstimo solicitado [SANTOS, 2000; SICREDI, 2003]. Caso contrário, as demais informações são coletadas para a análise do risco total de crédito.

GITMAN (1997) também cita o cinco C’s do crédito enfatizando os aspectos pessoais, caráter e capacidade, considerando-os como os fatores mais importantes para a concessão de empréstimos, pois:

*O analista de crédito geralmente dá maior importância aos dois primeiros C’s – caráter e capacidade – uma vez que eles representam os requisitos fundamentais para a concessão de crédito a um solicitante. A consideração para os demais C’s – capital, colateral e condições – é importante para a definição do acordo de crédito e tomada de decisão final, a qual depende da experiência e do julgamento do analista.*

Nesse sentido SOUZA *et al.* (2000) reporta-se ao caráter como sendo “... o histórico do solicitante quanto ao cumprimento de suas obrigações financeiras, contratuais e morais”.

Decidida a concessão de crédito ao cliente, as três questões básicas que servem para orientar a definição do limite de crédito são: [STEINER *et al.*, 1999, LIMA, 2002].

- Quanto o cliente merece de crédito?  
Depende da qualidade de risco apresentada, da capacidade de pagamento e do porte do cliente.
- Quanto pode ser oferecido ao cliente?  
Depende da capacidade de quem vai conceder o crédito.
- Quanto se deve, de fato, conceder de crédito ao cliente?  
Depende da política de crédito adotada.

### 2.3 – REDES NEURAIIS ARTIFICIAIS

As Redes Neurais Artificiais (RNA's) são uma maneira de criar modelos matemáticos, com capacidade de processar informações, inspirados em uma estrutura física natural: o cérebro humano. Estes modelos não pretendem replicar a operação do cérebro humano, apenas utilizam como inspiração fatores conhecidos sobre o seu funcionamento, ou seja, apresentando a capacidade de aprender, de tomar decisões, de se adaptar mesmo em presença de sinais ruidosos [HARTMANN, 2002]. São nestas características que se observa umas das maiores qualidades das RNA's: a capacidade de aprender através de exemplos e de generalizar sobre os padrões do conjunto de dados de treinamento, tornando possível a interpretação de outros padrões similares, mas não necessariamente idênticos aos utilizados durante o aprendizado.

HAYKIN (2001) define rede neural, como:

*“Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso”.*

Diversos tipos de redes neurais foram concebidos, podendo ser treinada através de diferentes algoritmos. Cada um destes tipos apresenta características diferentes, que variam de acordo com a forma como o mesmo é projetado [WONG e SELVI, 1998]. Para efeito deste estudo, destacam-se as redes neurais de múltiplas camadas alimentadas adiante (*feedforward*), treinadas de forma supervisionada através do algoritmo de Levenberg–Marquardt (LM), que pode ser considerado uma otimização do algoritmo *back-propagation*. O método de LM (também chamado de Método de Marquardt) trabalha muito bem na prática e tem-se tornado o método padrão a ser utilizado para resolver rotinas não-lineares de mínimo quadrados [LIMA, 2002]. O algoritmo LM é detalhado em BISHOP (1995).

### **2.3.1 – MÉTODO DE OTIMIZAÇÃO**

#### **2.3.1.1 – INTRODUÇÃO**

Um problema de otimização pode ser formulado como um problema de minimização (ou maximização) da seguinte forma: dada uma função  $f(x)$  representando um determinado sistema, sendo  $x$  o vetor de parâmetros desse sistema, o processo de otimização consiste em encontrar o valor mínimo (ou máximo) de  $f$ , ou mais genericamente o valor ótimo de  $f$ . Minimização e maximização são problemas relacionados, uma vez que encontrar o mínimo de  $f$  significa o mesmo que encontrar máximo de  $-f$ . Por isso, o termo “minimização” é comumente utilizado para referir-se também a problemas de maximização.

A solução verdadeira de um problema de minimização refere-se ao mínimo de todos os múltiplos mínimos, isto é, o mínimo global. Sua determinação é um problema difícil, devido à impossibilidade na maioria dos casos de saber se todos os mínimos foram encontrados. Algoritmos mais eficientes realizam diversas buscas

utilizando um esquema de minimização local, com cada busca partindo de um ponto diferente, aleatoriamente selecionado. A busca seria finalizada quando um nível aceitável de confiança de se ter encontrado o mínimo global fosse alcançado.

Seja uma função  $f(x)$  descrevendo um determinado sistema em termos de suas variáveis, sendo representadas na forma vetorial como (2.3.1).

$$x = [x_1 x_2 \cdots x_N]^T \quad (2.3.1)$$

O problema de minimização consiste em minimizar a função  $f$  de forma a obter seu valor mínimo, conforme (2.3.2).

$$\min_x F = f(x_{\min}) \quad (2.3.2)$$

Normalmente, o processo de minimização consiste em um processo iterativo, onde os valores dos parâmetros são ajustados até uma condição pré-determinada ser atingida. O valor ótimo final é indicado por  $F_{\min} = f(x_{\min})$ . Os ajustes dos parâmetros  $x$  são representados pelo vetor

$$\Delta x = [\Delta x_1 \Delta x_2 \cdots \Delta x_N]^T \quad (2.3.3)$$

o processo é iniciado a partir de uma solução inicial  $x_0$  correspondendo ao valor inicial  $F_0 = f(x_0)$  da função objetivo.

A solução obtida pode ser um mínimo local ou global. No primeiro caso, representa o menor valor de  $f$  em um intervalo ou espaço de busca finito. Já no segundo caso, representa o mínimo de  $f$  para qualquer variação de seus parâmetros  $x_i$ .

Alguns métodos de minimização requerem o uso do gradiente da função objetivo, obtido na forma das derivadas parciais de  $f$  com respeito aos parâmetros  $x_i$ . O uso dessas derivadas pode auxiliar na solução do problema de otimização; contudo, em certos casos seu uso pode não ser possível, devido à dificuldade de serem computadas as derivadas [SILVA, 2003].

As derivadas parciais de primeira ordem formam o vetor do gradiente jacobiano  $g$ , dado por

$$g = \nabla f = \left[ \frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \dots \quad \frac{\partial f}{\partial x_N} \right]^T \quad (2.3.4)$$

as derivadas parciais de segunda ordem, por sua vez, formam a matriz hessiana  $H$ , dada por

$$H = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 x_2} & \dots & \frac{\partial^2 f}{\partial x_1 x_N} \\ \frac{\partial^2 f}{\partial x_2 x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 x_N} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_N x_1} & \frac{\partial^2 f}{\partial x_N x_2} & \dots & \frac{\partial^2 f}{\partial x_N^2} \end{bmatrix} \quad (2.3.5)$$

A convergência pode ser verificada pelo progresso a cada iteração do valor de  $F$ . Pode-se considerar que a convergência tenha sido atingida quando  $F$  não reduz seu valor após um certo número de iterações. Todavia, pode não haver garantias que o mínimo global tenha sido alcançado. A velocidade de convergência é em geral associada ao número de iterações necessário para se obter a convergência, ou mais precisamente ao número de verificações do valor de  $F$ .

Métodos de otimização utilizando gradientes são baseados na expansão de Taylor [BISHOP, 1995]

$$f(x + \Delta x) = f(x) + g^T \Delta x + \frac{1}{2} \Delta x^T H \Delta x + \dots \quad (2.3.6)$$

os termos de terceira ordem e superiores são considerados desprezíveis. Caso os termos de segunda ordem sejam também desprezíveis, a equação anterior pode ser simplificada como

$$f(x + \Delta x) \approx f(x) + g^T \Delta x = F + \Delta F \quad (2.3.7)$$

e métodos baseados em gradientes são mais eficientes quando a função  $f$  tem derivadas contínuas  $C^2$  (derivadas de segunda ordem) que podem ser obtidas analiticamente. Em certos casos, porém, as derivadas somente podem ser obtidas por métodos numéricos.

Muitos métodos “globalmente convergentes”<sup>5</sup> adotam com critério de avaliação do progresso do algoritmo a seguinte função objetivo:

$$F(x) = \sum_{i=1}^N f_i^2 = \| \mathbf{f} \|^2 = \mathbf{f}^T \mathbf{f} \quad (2.3.8)$$

O método clássico para minimizar qualquer função não-linear  $F(x)$  é o método de Newton. Sendo  $x^{(l)}$  a  $l$ -ésima estimativa de  $F$ , o valor de  $F$  é dado por

$$F(x^{(l)} + \Delta x) = F(x^{(l)}) + \Delta x^T \nabla F(x^{(l)}) + \frac{1}{2} \Delta x^T \nabla^2 F(x^{(l)}) \Delta x + \dots \quad (2.3.9)$$

onde  $\nabla F = g^{(l)}$  e  $\nabla^2 F = H^{(l)}$  são o vetor gradiente e a matriz hessiana, contendo respectivamente as derivadas de primeira e segunda ordem de  $F(x)$ . A derivada de (2.3.9), desprezando os termos de terceira ordem e superiores, é dada por:

$$\nabla F(x) \approx g^{(l)} + H^{(l)}(x^{(l+1)} - x^{(l)}) \quad (2.3.10)$$

---

5 – O termo “globalmente convergente” significa: convergente para uma solução a partir de qualquer condição inicial

Considerando tão pequeno (quanto se queira) o  $\nabla F(x)$ , isto é,  $\nabla F(x) \approx 0$ , obtém-se:

$$H^{(l)}(x^{(l+1)} - x^{(l)}) = -g^{(l)} \quad (2.3.11)$$

Substituindo na equação anterior as expressões a seguir

$$H = 2 \left( A^T A + \sum_{i=1}^N \mathbf{f}_i \nabla^2 \mathbf{f}_i \right) \quad \text{e} \quad g = 2A^T \mathbf{f} \quad (2.3.12)$$

tem-se finalmente

$$(A^T A + B)^{(l)} u^{(l)} = -(A^T)^{(l)} \mathbf{f}^{(l)} \quad (2.3.13)$$

onde

$$B = \sum_{i=1}^N \mathbf{f}_i \nabla^2 \mathbf{f}_i \quad (2.3.14)$$

A matriz  $B$  pode ser interpretada como uma medida da não-linearidade das equações, sendo uma soma ponderada de suas derivadas de segunda ordem. Geralmente, considera-se que  $B$  seja desprezível em comparação com  $A^T A$ . Contudo, a ausência de  $B$  pode causar dificuldades no processo de minimização. Um modo de prover  $B$  é apresentado na seqüência através do Método de Levenberg – Marquardt.

### 2.3.1.2 – REDES DE MÚLTIPLAS CAMADAS ALIMENTADAS ADIANTE (*Feed-Foward Multilayer Perceptron*)

Tipicamente, a rede neural de múltiplas camadas consiste de um conjunto de unidades sensoriais (nós de fonte) que constituem as entradas da rede, uma ou

mais camadas ocultas de neurônios (unidades de processamento) e uma camada de saída de neurônios. As camadas estão ligadas por ligações sinápticas, cada uma caracterizada por um peso sináptico. O sinal de entrada se propaga para frente através da rede, camada por camada.

A Figura 2.3 apresenta uma rede neural do tipo *feed-forward* formada de uma camada de quatro nós de entrada, duas camadas ocultas de cinco e três neurônios, respectivamente, e uma camada com dois neurônios de saída.

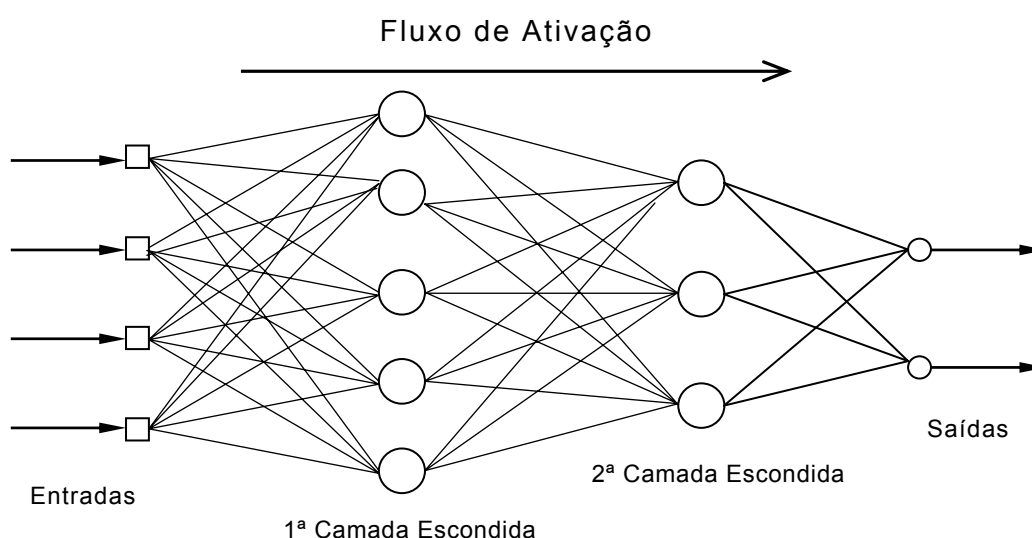


FIGURA 2.3 – Exemplo de Rede Neural Multicamadas *Feedforward*, totalmente conectada.  
 FONTE: DYMINSKI, 2000.

A Figura 2.4 mostra o modelo esquemático de um neurônio artificial que é a unidade de processamento da rede neural. As forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

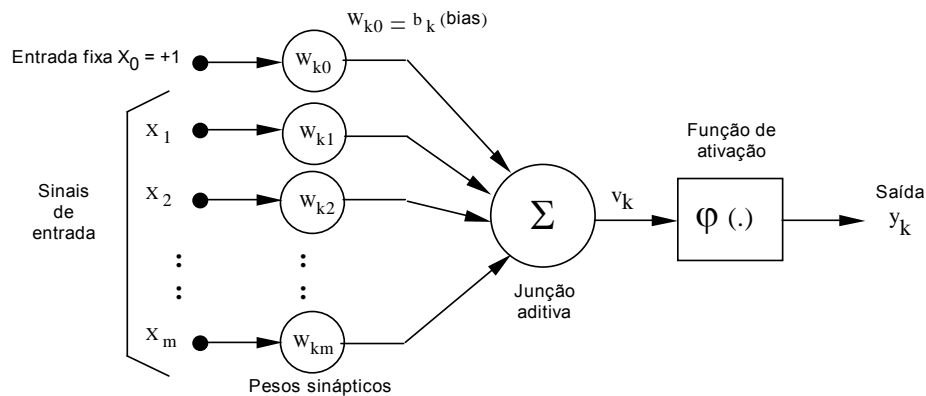


FIGURA 2.4 – Modelo não-linear de um neurônio

FONTE: HAYKIN, 2001.

Este é o modelo de neurônio proposto por MCCULLOCH – PITTS, onde são identificados três elementos básicos (HAYKIN, 2001):

- 1) Um conjunto de sinapses ou elos de conexão, cada uma caracterizada por um peso,  $\omega_{kj}$ , ou força própria;
- 2) Um “somador” para os sinais de entrada, ponderado pelas respectivas sinapses do neurônio (combinador linear) (Equação 2.3.15);

$$v_k = \sum_{j=0}^m \omega_{kj} x_j \quad (2.3.15)$$

- 3) Uma função de ativação para restringir a amplitude de saída de um neurônio, ou seja, limita o intervalo permissível de amplitude do sinal de saída a um valor finito. “[0,1] ou [-1,1]” (Equação 2.3.16).

$$y_k = \varphi(v_k) \quad (2.3.16)$$

Em termos matemáticos um neurônio  $k$  pode ser descrito através das equações (2.3.15) e (2.3.16), onde  $\varphi$  representa a função de ativação do neurônio artificial.

O *bias*  $b_k$  (parâmetro externo do neurônio artificial  $k$ ) tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação. Pode-se considerar o valor de *bias* no cálculo, introduzindo na Equação (2.3.15) uma nova sinapse de

índice 0. A sua entrada  $x_0$  assume o valor (+1) e o peso  $w_{k0}$  é igual a  $b_k$ .(equação 2.3.17)

$$v_k = \sum_{j=0}^m \omega_{kj} x_j + b_k \quad (2.3.17)$$

Destaca-se neste trabalho como função de ativação, a função sigmóide, que é definida como uma função estritamente crescente que exhibe um balanceamento adequado entre o comportamento linear e não-linear (MENON *et. al*, (1996), *apud* HAYKIN, 2001). A função de ativação existe para buscar um ajuste ótimo ou modelo para um problema. Conforme a Figura 2.5, são identificados três tipos básicos de funções de ativação: limiar, linear e sigmóide.

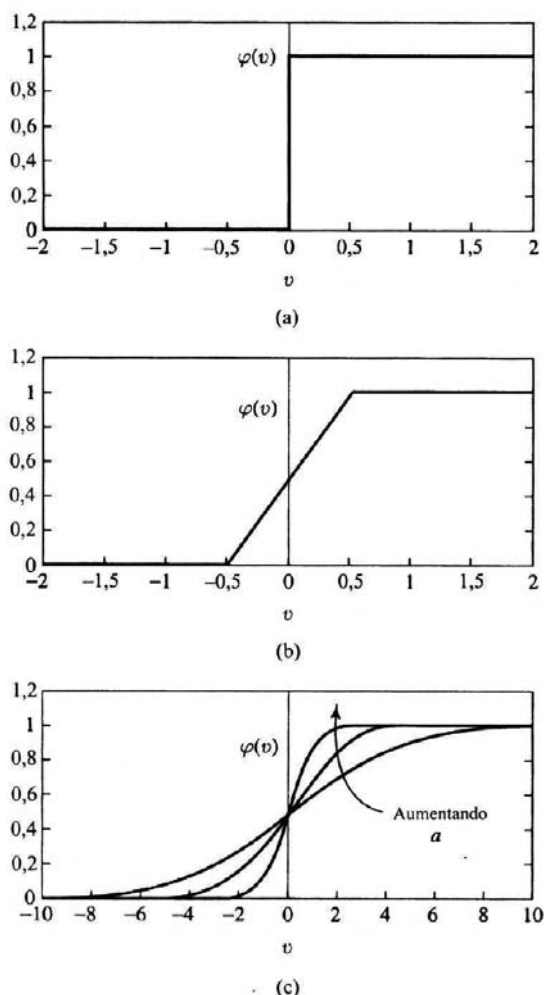


FIGURA 2.5 – (a) Função de limiar; (b) Função linear por partes; (c) Função sigmóide para parâmetro de inclinação  $\alpha$  variável

FONTE: HAYKIN, 2001.

Para BARRETO (2003), a função de mapeamento não-linear é dividida em duas partes: confluência e operação de ativação não-linear. A primeira é uma operação geral, a qual tem como argumentos os pesos sinápticos e entradas, sendo resolvido a operação pelo produto escalar; e a segunda é a função de ativação não-linear ( $\varphi$ ) que mapeia o valor de confluência para uma saída neural limitada.

Uma das maneiras mais conhecidas de se treinar uma rede neural é através do algoritmo de retro-propagação de erro, que é um algoritmo que utiliza o método do gradiente, ou seja, a direção de procura é conhecida a partir do gradiente, sendo possível especificar o passo dessa procura, definido por  $\eta$ , que é a velocidade de convergência do aprendizado, ou taxa de aprendizado.

BARRETO (2003) em seu trabalho com redes neurais baseou-se no algoritmo de retropropagação tradicional, contudo, a diferenciou em alguns aspectos: a otimização da camada intermediária foi auxiliada por Algoritmo Genético (AG) e, além disso, fez a incorporação de operadores lógicos AND/OR, ao invés do somatório de pesos utilizando formalização da rede lógica *fuzzy*.

Do algoritmo de retropropagação do erro, um dos muitos importantes aspectos é a eficiência computacional. O treinamento envolve um procedimento iterativo para minimizar a função de erro, sendo ajustados os pesos em uma seqüência de dois passos [BISHOP, 1995]. No primeiro passo, as derivadas da função erro com respeito aos pesos precisam ser avaliadas; no segundo passo, as derivadas são então utilizadas para ajustar os pesos.

Calculam-se, em primeiro lugar, os erros das saídas e só depois é que o algoritmo vai “retropropagando” o valor do erro calculado para as camadas anteriores. O sinal de erro estimado na saída do neurônio  $k$ , na  $n$ -iteração, é definido pela equação (2.3.18):

$$e_k(n) = d_k(n) - y_k(n) \quad (2.3.18)$$

Em HAYKIN (2001), se o neurônio  $k$  pertencer a camada de saída, o sinal de erro  $e_k(n)$  tem o propósito de aplicar uma seqüência de ajustes corretivos aos pesos sinápticos  $w_{kj}(n)$ ; onde  $w_{kj}(n)$  representa o peso sináptico conectando a saída do neurônio  $j$  à entrada do neurônio  $k$ , na iteração  $n$ . Estes ajustes tendem a aproximar o sinal de saída  $y_k$  da resposta desejada  $d_k$ . Este objetivo é alcançado minimizando uma função de custo,  $\varepsilon(n)$ , que é o valor instantâneo da energia do erro para o neurônio  $k$  definido pela equação (2.3.19):

$$\varepsilon(n) = \frac{1}{2} [e_k(n)]^2 \quad (2.3.19)$$

O valor instantâneo,  $\varepsilon(n)$ , da energia total do erro é obtido somando-se os termos de todos os neurônios da camada de saída:

$$\varepsilon(n) = \frac{1}{2} \sum_{k \in C} [e_k(n)]^2 \quad (2.3.20)$$

onde o conjunto  $C$  inclui todos os neurônios da camada de saída da rede.

Sendo determinado  $e_k(n)$ , calcula-se diretamente o gradiente local  $\delta_k(n)$  para o neurônio de saída  $k$ , que é definido pela equação (2.3.21).

$$\delta_k(n) = -\frac{\partial \varepsilon(n)}{\partial v_k(n)} \quad , \text{ ou seja,}$$

$$\delta_k(n) = e_k(n) \varphi'_k(v_k(n)) \quad (2.3.21)$$

Efetuada as devidas relações de derivação para o algoritmo de retropropagação, tem-se que a correção de peso ( $\Delta w_{kj}(n)$ ) aplicada ao peso sináptico conectando o neurônio  $j$  ao neurônio  $k$  é definida pela regra delta:

$$\Delta\omega_{kj}(n) = \eta\delta_k(n)y_j(n) \quad (2.3.22)$$

onde:

$\eta$ : parâmetro da taxa de aprendizagem;

$\delta_k(n)$ : gradiente local;

$y_j(n)$ : sinal de entrada do neurônio  $k$ .

Para o cálculo da correção de peso, pelo método de descida mais íngreme, tem-se uma “aproximação” para a trajetória dos pesos. Quanto menor for  $\eta$ , mais suave será a trajetória no espaço de pesos; porém, esta melhoria é obtida às custas de uma taxa de aprendizagem lenta. Entretanto, se for adotado  $\eta$  muito grande, para acelerar a taxa de aprendizagem, a rede pode se tornar instável (oscilante).

Como mostrado por RUMELHART *et al.* (1986a), *apud* HAYKIN (2001), um método para evitar o perigo da instabilidade é aumentar a taxa de aprendizagem, simultaneamente, incluindo um termo de “momentum”. Logo, tem-se como correção de peso:

$$\Delta\omega_{kj}(n) = \alpha\Delta\omega_{kj}(n-1) + \eta\delta_k(n)y_j(n) \quad (2.3.23)$$

HAYKIN (2001) salienta que para a retropropagação quando o neurônio  $j$  está localizado em uma camada oculta da rede, não existe uma resposta desejada especificada para aquele neurônio. Daí a necessidade do cálculo para o gradiente local  $\delta_j(n)$ :

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n)\omega_{kj}(n), \text{ o neurônio } j \text{ é oculto} \quad (2.3.24)$$

No passo para frente, os pesos sinápticos se mantêm inalterados em toda a rede e os sinais funcionais da rede são calculados individualmente, neurônio por

neurônio. Assim, a fase de propagação começa na primeira camada oculta, com a apresentação do vetor de entrada, e termina na camada de saída calculando o sinal de erro de cada neurônio desta camada. Por outro lado, o passo de retropropagação começa na camada de saída passando-se os sinais de erro através da rede, camada por camada, e recursivamente calculando-se o  $\delta$ , gradiente local, de cada neurônio. Este processo recursivo permite que os pesos sinápticos sofram modificações de acordo com a regra delta da Equação (2.3.22).

A Figura 2.6 mostra uma representação em grafo de fluxo de sinal do processo de aprendizagem por correção de erro, enfocando a atividade na vizinhança do neurônio  $k$ .

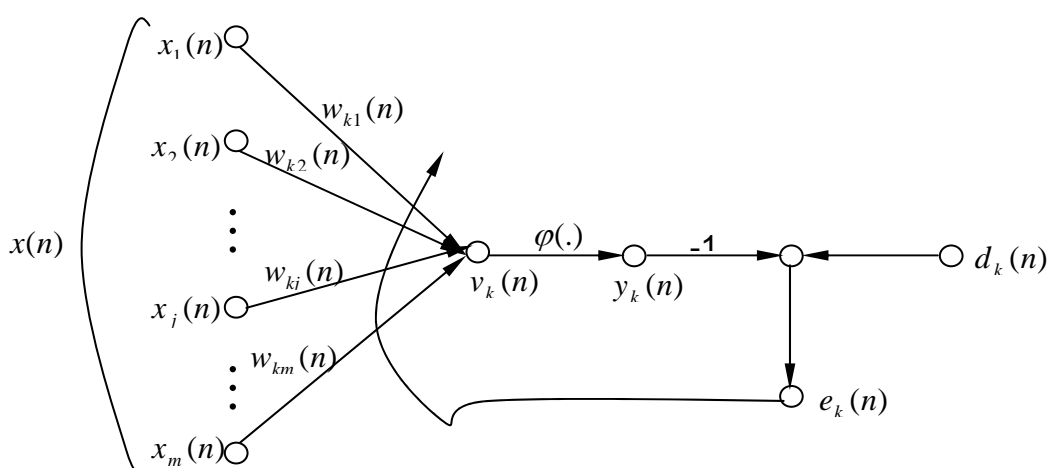


FIGURA 2.6 – Ilustração da aprendizagem por correção de erro para o neurônio  $k$ .  
FONTE: HAYKIN, 2001.

### 2.3.1.3. – MÉTODO DE LEVENBERG–MARQUARDT

A principal limitação do método de Newton-Raphson, ou simplesmente método de Newton, envolve  $A$ , a matriz jacobiana. A matriz  $A^T A$ , que precisa ser

invertida a cada iteração, pode apresentar problemas de mal-condicionamento, o que pode conduzir a resultados divergentes, e até mesmo à singularidade de  $A^T A$ . É, portanto, essencial adotar uma estratégia alternativa quando ocorre um mal-condicionamento extremo. Uma aproximação usual é o uso do método de Levenberg–Marquardt [SILVA, 2003]: modifica-se a matriz a inverter de forma a evitar o mal-condicionamento, fazendo-se  $B = \mu I$  na equação (2.3.13) ( $\mu > 0$  e  $I$  é a matriz identidade), resultando em

$$(A^T A + \mu I)^{(l)} u^{(l)} = -(A^T)^{(l)} f^{(l)} \quad (2.3.25)$$

esta modificação altera apenas as componentes  $a_{ii}$  da diagonal principal de  $A^T A$ , resultando em componentes modificadas  $a'_{ii}$  de forma que  $a'_{ii} = a_{ii} + \mu$ . Algumas implementações realizam a modificação  $a'_{ii} = a_{ii}(1 + \mu)$ , porém com efeitos similares.

Deve-se ter um cuidado maior na escolha do valor do parâmetro  $\mu$ . Se o valor de  $\mu$  for muito grande, a matriz  $\mu I$  torna-se dominante em relação a  $A^T A$ , a direção de busca tende à direção do máximo declive (*steepest-descent* – desenvolvido por Cauchy em 1847) [LIMA, 2002; SILVA, 2003] e o módulo do vetor de busca torna-se pequeno. Por outro lado, se o valor de  $\mu$  for muito pequeno, as direções de busca são similares às do método de Newton-Raphson. Deve-se, então, escolher um valor para  $\mu$  de forma a evitar essas duas situações extremas. Este valor precisa ser relacionado com a magnitude (módulo) de  $A^T A$ , uma medida provida pelo traço dessa matriz:  $tr(A^T A) = \sum \lambda_i = \lambda(A^T A)_{ii}$  (a soma dos auto-valores, ou a soma dos elementos da diagonal principal da matriz). Tipicamente,  $\mu$  apresenta valores da ordem de  $10^{-4}$  [LIMA, 2002].

O valor do parâmetro  $\mu$  pode ser alterado a cada iteração, de forma a obter uma redução do valor da função objetivo  $F$ , ou seja: em cada iteração  $l$ ,  $F^{(l)}$  é avaliado. Caso  $F^{(l)} < F^{(l-1)}$  o valor de  $\mu_l$  é reduzido, e passa-se para a próxima iteração. Contudo, caso  $F^{(l)} > F^{(l-1)}$  o valor de  $\mu_l$  é aumentado, e  $F^{(l)}$  é novamente

avaliado. O processo é repetido até a condição  $F^{(l+1)} < F^{(l)}$  seja enfim satisfeita. O Quadro 2.3 resume este método.

- 1) Selecionar uma solução inicial  $x^{(0)}$  e dois critérios de término  $\varepsilon_1$  e  $\varepsilon_2$ . Calcular  $A^{(0)}$ ,  $f^{(0)}$  e  $g^{(0)}$ . Se  $\|f^{(0)}\| \leq \varepsilon_1$  ou  $\|g^{(0)}\| \leq \varepsilon_2$  então terminar o algoritmo; senão fazer  $l=0$  e avançar para o segundo passo.
- 2) Calcular  $A^{(l)}u^{(l)} = -f^{(l)}$ . Se  $A^{(l)}$  for singular então calcular  $(A^T A)^{(l)}$ ,  $\tau = \text{tr}(A^T A)^{(l)}$  e  $(A^T A + \mu I)^{(l)}$ . Calcular  $(A^T A + \mu \tau I)^{(l)} u^{(l)} = -f^{(l)}$ .
- 3) Realizar uma busca ao longo da direção  $u^{(k)}$  de forma a determinar um passo  $\alpha_l$  satisfazendo à regra de Armijo [SILVA, 2003].
- 4) Calcular  $x^{(l+1)} = x^{(l)} + \alpha_l u^{(l)}$ .
- 5) Calcular  $f^{(l+1)}$ . Se  $\|f^{(l+1)}\| \leq \varepsilon_1$  então terminar o algoritmo; senão calcular  $A^{(l+1)}$  e  $g^{(l)}$ . Se  $\|g^{(l+1)}\| \leq \varepsilon_2$  então terminar o algoritmo; senão fazer  $l = l + 1$  e voltar para o segundo passo.

QUADRO 2.3 – Algoritmo de Levenberg–Marquardt

Enquanto o *back-propagation* padrão utiliza a descida de gradiente como método de aproximação do mínimo da função erro, o algoritmo de Levenberg–Marquardt utiliza uma aproximação pelo método de Newton.

Têm-se discutido algoritmos de otimização com o intuito de um melhor projeto para as redes neurais com uma ampla variedade da função erro. Uma

comparação do algoritmo LM com outros métodos de treinamento, MLP, é destacado em BISHOP, 1995.

### 2.3.2 – SEPARABILIDADE DE PADRÕES

A compreensão sobre a operação de uma rede neural como um classificador de padrões, pode ser desenvolvida estudando a separabilidade de padrões (HAYKIN, 2001).

Seja uma família de superfícies em que cada uma divide o espaço de entradas em duas regiões. Seja  $X$  um conjunto de  $N$  vetores de entrada,  $x_1, x_2, x_3, \dots, x_N$ , em que cada um é atribuído a uma de duas classes  $X_1$  ou  $X_2$ . Esta partição binária é separável, segundo a família de superfícies, se existir uma superfície na família que separe os pontos da classe  $X_1$  dos da classe  $X_2$ . Para cada padrão de entrada  $x \in X$ , define-se um vetor composto de funções reais:

$$\varphi(x) = [\varphi_1(x), \varphi_2(x), \dots, \varphi_{m_1}(x)]^T \quad (2.3.26)$$

onde,  $\varphi_i(x)$  é denominado função oculta. Suponha que  $x$  é um vetor de dimensão  $m_0$ . O vetor  $\varphi(x)$  mapeia pontos de um espaço  $m_0$  dimensional em pontos de um espaço  $m_1$  dimensional. O espaço definido pelo conjunto das funções ocultas,  $\{\varphi_i(x)\}_{i=1}^{m_1}$  denota-se como sendo o espaço oculto ou espaço de características.

Segundo COVER (1965), *apud* HAYKIN (2001), uma partição binária  $\{X_1, X_2\}$  de  $X$  é  $\varphi$ -separável se existir um vetor  $m_1$ -dimensional  $w$  tal que:

$$w^T \varphi(x) > 0, \quad x \in X_1 \quad (2.3.27)$$

$$w^T \varphi(x) < 0, \quad x \in X_2$$

E, nesse caso,

$$x : w^T \varphi(x) = 0 \quad (2.3.28)$$

define a superfície de separação no espaço de entrada, correspondente ao hiperplano separador no espaço  $\varphi$ .

Em resumo, segundo HAYKIN (2001), sobre o Teorema de COVER (1965), há duas características fundamentais para a separabilidade de padrões:

1. Funções ocultas,  $\varphi_i(x)$ , não lineares, onde  $x$  é o vetor de entrada para  $i = 1, 2, 3, \dots, m_1$ ;
2. Grande dimensionalidade do espaço oculto,  $m_1$ , comparado com o espaço de entrada.

Salienta-se que a má aprendizagem pode ser decorrente de algum dos seguintes fatores:

- não existe uma saída definida para cada entrada, tendo como consequência falha para sua existência;
- falta de informação no conjunto de treino para construir a hipersuperfície de classificação, ocorrendo falha para a unicidade;
- demasiado ruído no conjunto de treino que pode dar origem a uma saída não pertencente a  $Y$  para um  $x \in X$ , ou seja, ocorre descontinuidade.

## **2.4 – ANÁLISE MULTIVARIADA**

### **2.4.1 – ANÁLISE DE COMPONENTES PRINCIPAIS**

A análise de componentes principais (ACP) consiste essencialmente em reescrever as coordenadas das amostras em outro sistema de eixo mais conveniente para a análise dos dados [JOHNSON and WICHERN, 1998]. Em outras palavras, aplica-se uma transformação ortogonal às n-variáveis originais obtendo-se as suas combinações lineares que são chamadas de componentes principais. A principal característica das componentes principais, além da ortogonalidade, é que a sua importância na estrutura de covariância do vetor aleatório observado é proporcional à magnitude do autovalor correspondente. Torna-se evidente, então, que a transformação ortogonal é obtida com base nos autovalores e autovetores da matriz de covariância (ou de correlação) do vetor aleatório observado. Então, elas podem ser colocadas em ordem decrescente de importância e nas direções de máxima variância, ou seja, a componente principal 1 detém mais informação estatística que a componente principal 2, que por sua vez tem mais informação estatística que a componente principal 3 e assim por diante.

### **2.4.2 – ANÁLISE FATORIAL**

A Análise Fatorial (AF) teve seu início mais recente, no começo do século XX, através de Karl Pearson, Charles Spearman e outros, sendo definida como uma melhor medida “inteligente” na redução de dados [CHATFIELD e COLLINS, 1992; JOHNSON e WICHERN, 1998]. Exemplos da Análise Fatorial em marketing usando dados com escalas ordinais podem ser encontrados em JOHNSON e WICHERN (1998) que dizem respeito ao estudo da preferência do consumidor sobre os atributos de um novo produto, através da pontuação, numa escala de diferencial semântico de 7 pontos.

Para JOHNSON e WICHERN (1998) a Análise Fatorial é uma técnica de análise multivariada e pode ser considerada como uma extensão da ACP. Existe consenso comum entre os autores sobre a importância e a aplicação da AF que objetiva examinar a interdependência entre variáveis, sendo sua principal característica a capacidade de redução de dados, isto é, resumir um grande número de variáveis originais, num pequeno número de variáveis com entendimento dos dados.

A Análise Fatorial possui, basicamente, quatro etapas para a sua elaboração que são (LATIF, 2000):

- 1) Cálculo da matriz de correlação das variáveis em estudo para a verificação do grau de associação entre as variáveis, duas a duas. Nesta etapa, é verificada a adequação da aplicação da AF;
- 2) Extração dos fatores mais significativos que representarão os dados. Aqui se sabe quão bem o método representa os dados;
- 3) Aplicação de rotação, nos fatores, para facilitar o entendimento dos mesmos;
- 4) Geração dos escores fatoriais para utilização em outras análises.

#### **2.4.2.1 – MODELO FATORIAL ORTOGONAL**

Seja o vetor aleatório  $X$ , com distribuição  $X \sim .(\mu, \Sigma)$ , com  $p$  componentes, tendo média  $\mu$  e matriz de covariância  $\Sigma$ . No modelo fatorial  $X$  é linearmente dependente, sobre algumas variáveis aleatórias não-observáveis e  $F_1, F_2, \dots, F_m$

(para  $m \leq p$ ) chamadas fatores comuns e  $p$  fontes de variações aditivas sendo  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$  chamados erros ou fatores específicos.

Em particular JOHNSON e WICHERN (1998) destacam como modelo de Análise Fatorial, o seguinte formato:

$$\begin{aligned}
 X_1 - \mu_1 &= \ell_{11}F_1 + \ell_{12}F_2 + \dots + \ell_{1m}F_m + \varepsilon_1 \\
 X_2 - \mu_2 &= \ell_{21}F_1 + \ell_{22}F_2 + \dots + \ell_{2m}F_m + \varepsilon_2 \\
 &\dots \qquad \dots \qquad \dots \\
 X_p - \mu_p &= \ell_{p1}F_1 + \ell_{p2}F_2 + \dots + \ell_{pm}F_m + \varepsilon_p
 \end{aligned} \tag{2.4.1}$$

onde:

$\mu_i$  = média da  $i$ -ésima variável;

$\varepsilon_i$  =  $i$ -ésimo erro ou fator específico;

$F_j$  =  $j$ -ésimo fator comum;

$\ell_{ij}$  = peso ou carregamento na  $i$ -ésima variável  $X_i$  do  $j$ -ésimo fator  $F_j$ .

$i = 1, 2, \dots, p$  e  $j = 1, 2, \dots, m$

ou, em notação matricial:

$$\underset{(px1)}{X} - \underset{(px1)}{\mu} = \underset{(pxm)}{L} \cdot \underset{(mx1)}{F} + \underset{(px1)}{\varepsilon} \tag{2.4.2}$$

Assumindo que:

$$E(F) = \underset{(mx1)}{0}, \qquad Cov(F) = E[FF'] = \underset{(mxm)}{I}$$

onde a estrutura da covariância, "Cov", para o modelo fatorial ortogonal é (JOHNSON e WICHERN, 1998):

$$\begin{aligned}
 1. \text{Cov}(X) &= LL' + \psi & \text{ou,} & & \text{Var}(X_i) &= l_{i1}^2 + \dots + l_{im}^2 + \psi_i \\
 & & & & \text{Cov}(X_i, X_k) &= l_{i1}l_{k1} + \dots + l_{im}l_{km} \\
 2. \text{Cov}(X, F) &= L & \text{ou} & & \text{Cov}(X_i, F_j) &= l_{ij}
 \end{aligned}$$

Logo,

$$E(\varepsilon) = \underset{(p \times 1)}{0}, \quad \text{Cov}(\varepsilon) = E[\varepsilon\varepsilon'] = \underset{(p \times p)}{\Psi} = \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_p \end{bmatrix} \quad (2.4.3)$$

sendo que  $\psi_1, \psi_2, \dots, \psi_p$ , são as variâncias de  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ , respectivamente; e  $F$  e  $\varepsilon$  são independentes, ou seja,

$$\text{Cov}(\varepsilon, F) = E(\varepsilon F') = \underset{(p \times m)}{0} \quad (2.4.4)$$

Essas hipóteses e o modelo vistos constituem o Modelo Fatorial Ortogonal.

## CAPÍTULO III

### 3.0 – MATERIAIS E MÉTODOS

#### 3.1 – INTRODUÇÃO

As cooperativas, como lembra o Juiz Federal Renato Lopes Becho, em Legislação e formalística de Crédito (SICREDI, 2003):

*“São associações de pessoas que se unem para atingir determinado fim específico, organizando um empreendimento com características diversas das demais pessoas jurídicas, tanto civis quanto comerciais”.*

Muito embora as atividades desenvolvidas pelas cooperativas de crédito sejam semelhantes às praticadas por bancos, existem inúmeras diferenças entre as duas instituições. O Quadro 3.1 relaciona as principais distinções entre banco e cooperativa de crédito.

BANCO	COOPERATIVA DE CRÉDITO
<ul style="list-style-type: none"> <li>✓ Busca lucratividade própria;</li> <li>✓ Captação e aplicação livre no mercado;</li> <li>✓ Pluralidade de clientes / correntistas;</li> <li>✓ Há descompromisso com os clientes no que se refere à assistência mútua e o proveito comum.</li> </ul>	<ul style="list-style-type: none"> <li>✓ Ausência do objetivo do lucro próprio;</li> <li>✓ Mutualidade e assistência entre os associados;</li> <li>✓ Restrição aos correntistas – somente associados pode operar com a cooperativa de crédito;</li> <li>✓ Operações com taxas mais baixas que o mercado.</li> </ul>

QUADRO 3.1 – Principais distinções entre Banco e Cooperativa de Crédito.

FONTE: SICREDI (2003)

Não fossem todas estas robustas diferenças, a Lei nº 5.764/71 dispõe no Parágrafo Único, do Artigo 5º:

*“É vedado às cooperativas o uso da expressão ‘BANCO’”*

### **3.2 – BASE DE DADOS**

O processo decisório de crédito para pessoas físicas, clientes da cooperativa, é baseado em dados e fatos observados do proponente, tais como: profissão; idade; sexo; estado civil; saldo da conta capital; cartão de crédito; média mensal da conta corrente; média trimestral da conta corrente; histórico dos últimos doze meses, quanto a pontualidade das operações contratadas referentes ao crédito; quantidade de cheques devolvidos; existência de títulos de cobrança e outros.

O grau de risco de cada operação é apurado com a classificação de risco do tomador e suas características e, também, da avaliação do grau de risco da operação, conforme o estabelecido na Resolução nº 2.682, de 21-12-99, do Conselho Monetário Nacional [SICREDI, 2003]. A classificação adotada por essa resolução definiu as classes de risco e o respectivo provisionamento para as operações de crédito abrigadas em cada classe. Assim, o Banco Central do Brasil (BACEN) determina a classificação de seus clientes em diversas categorias. Para cada grau de risco é estabelecido um nível de provisionamento adequado à probabilidade de perda, conforme o Quadro 3.2. Assim, as operações são classificadas em 9 níveis de risco, sendo “AA” o menor grau de risco e “H” o maior grau de risco, e, segundo o Manual de Políticas de Crédito do SICREDI, a ocorrência de restrições junto ao SPC, Serviço de Proteção ao Crédito, ou SERASA acarreta no rebaixamento do grau de risco do associado/cliente para risco “H”.

CLASSIFICAÇÃO DE RISCO DA OPERAÇÃO	% A PROVISIONAR
AA	0%
A	0,5%
B	1%
C	3%
D	10%
E	30%
F	50%
G	70%
H	100%

QUADRO 3.2 – Classificação de Risco.

FONTE: SICREDI (2003)

O enquadramento das pessoas físicas nos diversos níveis obedece a vários critérios de avaliação, como situação sócio-econômico-financeira; nível de endividamento; capacidade de geração de resultados futuros, pontualidade de pagamento; setor de atividade econômica; renda e despesas; patrimônio e registro de pendências. Os classificados como F, G e H certamente ficarão de fora da carteira de empréstimos do sistema financeiro.

Ainda, segundo o Manual de Políticas de Crédito do SICREDI, serão formalizadas novas operações de crédito exclusivamente às que apresentem classificação “D” ou melhor. A mesma resolução determina que as operações sejam revisadas mensalmente e com base nos dias de atrasos é imposta nova classificação de risco. O Quadro 3.3 mostra a relação entre o atraso e o nível de risco.

Níveis de risco	A	B	C	D	E	F	G	H
Dias de atraso	Até 14 dias	De 15 a 30 dias	De 31 a 60 dias	De 61 a 90 dias	De 91 a 120 dias	De 121 a 150 dias	De 151 a 180 dias	Acima de 180 dias

QUADRO 3.3 – Classificação quanto ao atraso no pagamento.

FONTE: SICREDI (2003)

Assim, atrasar o pagamento de crédito em mais de 15 dias, custará o rebaixamento de um cliente AA para o nível A, por exemplo.

Para o desenvolvimento dos modelos propostos neste trabalho, houve a necessidade de se construir uma base de dados usando informações fornecidas pela Cooperativa de Crédito. Para realizar esta tarefa, o método utilizado baseou-se em um processamento efetuado em fases onde cada uma das fases se torna geradora de entradas para a fase seguinte.

É importante salientar que a eficiência de uma técnica é proporcional à qualidade das informações disponíveis. Assim, a fase de coleta de dados destaca-se de suma importância. A primeira fase foi a identificação no cadastro utilizado para o processamento, das informações disponíveis que teriam relevância ao estudo. Embasado nos C's do crédito, descrito no capítulo II, seguiu-se para a coleta dos dados disponíveis.

Na garimpagem dos dados, foram utilizados somente os clientes que dispunham de todas informações em comum, isto é, cooperado com informação incompleta, automaticamente foi eliminado da base de dados, ou seja, não entrou nas análises seguintes. No *"layout"* do arquivo-base foram coletadas informações de 1.017 cooperados com 23 campos (variáveis do modelo). Na seqüência, procedeu-se ao descarte de variáveis de acordo com o descrito no capítulo II. O descarte diminui o custo de levantamento dos dados futuros.

Dentre os campos eliminados, estão os dados de identificação do proponente: nome, endereço, sexo, CPF, dados do cônjuge e outros. Também foram desprezados campos com grande número de repetições de uma atribuição específica (ex. "profissão: agricultor" ou "utiliza cartão de crédito: não"). A atividade exercida pelo proponente tem grande importância sobre o grau de risco futuro do mesmo, em função do comportamento de cada mercado, havendo mercados tradicionalmente mais arriscados e mercados mais estáveis [SICREDI, 2003], e devido ao grande número de repetições atribuídas a estas variáveis (superiores a 92%) foi efetuado o descarte das mesmas.

Referente ao perfil dos associados foram selecionados 9 campos (variáveis), sendo que o critério de representatividade do campo também foi utilizado. Uma visão detalhada das variáveis pode ser observada no Quadro 3.4, sendo que as variáveis 2, 3, 4, 5, 7 e 8 foram normalizadas entre 0 e 1. Já as variáveis 1, 6 e 9, que são variáveis não-numéricas, tiveram que ser quantificadas.

DESCRIÇÃO		CONDIÇÕES	VALORES	SIGNIFICADO
01	Informações excludentes	S P C / SERASA	0 1	Não Sim
02	Limite de crédito	CHEQUE ESPECIAL	0 1	Mínimo Máximo
03	Média trimestral	CONTA CORRENTE	0 1	Mínimo Máximo
04	Conta capital	SALDO	0 1	Mínimo Máximo
05	Cheques devolvidos	ALÍNEA (11, 12, 13) <sup>6</sup>	0 1	Mínimo Máximo
06	Cobrança	EXISTÊNCIA	0 1	Não Sim
07	Histórico dos últimos 12 Meses, referente a crédito.	NÚMERO DE PARCELAS PAGAS	0 1	Mínimo Máximo
08		PARCELAS PAGAS COM ATRASO A PARTIR DE 15 DIAS	0 1	Mínimo Máximo
09	Inadimplência	EXISTEM OPERAÇÕES DE CRÉDITO EM INADIMPLÊNCIA PARA O ASSOCIADO?	0 1	Não Sim

QUADRO 3.4 – Campos selecionados e valores adotados.

6 – Alíneas das devoluções de cheques:

Alínea 11, primeira apresentação; Alínea 12, segunda apresentação (inclui no serasa); Alínea 13, conta encerrada (inclui no serasa).

A seleção inicial pela busca das variáveis, partiu do cadastro do qual foram selecionados os registros com as operações que estão classificadas em 9 níveis de risco, sendo “AA” o menor grau e “H” o maior grau de risco. É observado no manual de políticas de crédito da cooperativa que novas operações de crédito ocorrerão exclusivamente aos clientes que apresentarem classificação “D” ou melhor. Desta forma optou-se em utilizar “1” aos cooperados que deveriam receber crédito e “0” caso contrário. O conjunto total da amostra para tratamento dos dados ficou composto de 961 observações(cooperados) com 9 variáveis cada, sendo estas detalhadas no Quadro 3.4.

O conjunto de dados disponível ao processamento foi dividido aleatoriamente, em conjuntos para treinamento, para validação e para teste, conforme o Quadro 3.5, de forma que os mesmos não se repetissem (conjuntos exclusivos), nem ficassem saturados com determinadas informações, como a ocorrência de restrições junto ao SPC ou SERASA que acarreta no rebaixamento do grau de risco do cooperado para risco “H” onde há a implicação da imediata recusa ao fornecimento do crédito ao cliente.

Com o objetivo de resguardar o sigilo, por força legal, todas as contas foram descaracterizadas, omitindo nomes, endereços, datas de nascimento, CPF, RG, etc. Os números das contas foram substituídos por um número de controle, gerado aleatoriamente, de modo a garantir que cada conta tenha um número único. Posteriormente, a tabela original foi descartada de modo a tornar impossível a identificação do número original da conta ou o titular. Nesta fase, os arquivos originais também foram descartados. O arquivo final, com 9 variáveis referentes a 961 registros (observações), foi base para todos os processos seguintes, gerando uma matriz de 961 linhas por 9 colunas.

O número total de cooperados analisados foi formatado a partir do arquivo recebido, em um “*layout*” padrão de modo a atender os percentuais do Quadro 3.5, sendo estes dados cadastrais os disponíveis ao processamento.

VALORES CORRESPONDENTES PARA REDE NEURAL		
Número total de cooperados analisados: 961		
TREINAMENTO: 577 cooperados (60,04% do total)	VALIDAÇÃO: 192 cooperados (19,68% do total)	TESTE: 192 cooperados (19,68% do total)
REFERENTE AO FORNECIMENTO DE CRÉDITO	REFERENTE AO FORNECIMENTO DE CRÉDITO	REFERENTE AO FORNECIMENTO DE CRÉDITO
Fornece: 67,07% não fornece: 32,06% SPC/SERASA: 0,87%	Fornece: 64,59% não fornece: 34,37% SPC/SERASA: 01,04%	Fornece: 61,46% Não fornece: 36,98% SPC/SERASA: 01,56%

QUADRO 3.5 – Percentuais referentes às observações.

A ênfase ao fornecimento de crédito, ou não, na fase de treinamento da rede, é destacado na primeira coluna do Quadro 3.5, onde o percentual de 67,07% se refere ao número de cooperados que são merecedores de crédito, ou seja, se enquadraram segundo a classificação nível “D” ou melhor. E, em contrapartida, o percentual de 32,06% diz respeito ao não fornecimento de crédito ao cliente devido a sua classificação. Já o percentual de 0,87% se refere ao número de clientes que se encontram junto ao SPC ou SERASA. É ressaltado que tanto os percentuais da coluna referente ao treinamento como validação e teste da rede, dizem respeito ao subtotal de cada subconjunto e não ao total geral de cooperados.

### 3.3 – PROCEDIMENTOS DE PREPARAÇÃO DOS DADOS

Nesta pesquisa, além do STATGRAPHICS Plus 5.1, também foi utilizado o pacote computacional MATLAB 6.5 (*release* 13), onde foram construídas funções,

com o intuito de uma melhor performance da rede. O equipamento utilizado foi um computador Pentium 4, CPU 2.40 GHz, com 480 mega bytes de memória (RAM).

Para uma avaliação estatística das características dos dados, foi utilizada a ACP que teve o propósito de avaliar a importância relativa das diferentes variáveis envolvidas no processo da análise de crédito. Este procedimento pode ser adotado em algumas situações em que o número de entradas do modelo é grande, porém as mesmas encontram-se altamente correlacionadas. Esta análise pode ser útil nestes casos, pois possibilita a redução da dimensão dos vetores de entrada.

A técnica da Análise de Componentes Principais aplicada na base de dados mostrou que um modelo com seis componentes principais seria suficiente para representar a estrutura de covariância inicial, com 10% de perda de informação. Na Tabela 3.1 observa-se a proporção da variância explicada pelos autovalores ordenados de forma decrescente.

ORDEM	AUTOVALORES	VAR. EXPL. (em %)	VAR. EXPL. ACUM. (em %)
1	2.8654	31.84	31.84
2	1.4915	16.57	48.41
3	1.0491	11.66	60.07
4	0.9546	10.61	70.67
5	0.9133	10.15	80.82
6	0.8271	09.19	90.01
7	0.3963	04.40	94,41
8	0.2921	03.25	97,66
9	0.2107	02.34	100.00

TABELA 3.1 – Autovalores em ordem decrescente de magnitude.

A ACP também pode ser usada para julgar a importância das próprias variáveis originais escolhidas, ou seja, as variáveis originais com maior peso na combinação linear dos primeiros componentes principais são as mais importantes do ponto de vista estatístico e, como o objetivo principal é realizar boa compressão de

dados preservando o máximo possível de informação sobre as entradas, o uso da ACP oferece um procedimento útil de aprendizagem auto-organizada.

A Figura 3.1 indica que, a partir do sexto fator os autovalores apresentam uma disposição levemente mais próxima da parte inferior do gráfico e, aproximadamente, da forma horizontal. Este pode ser considerado o ponto de corte do número de fatores a serem utilizados no modelo que está sendo construído.

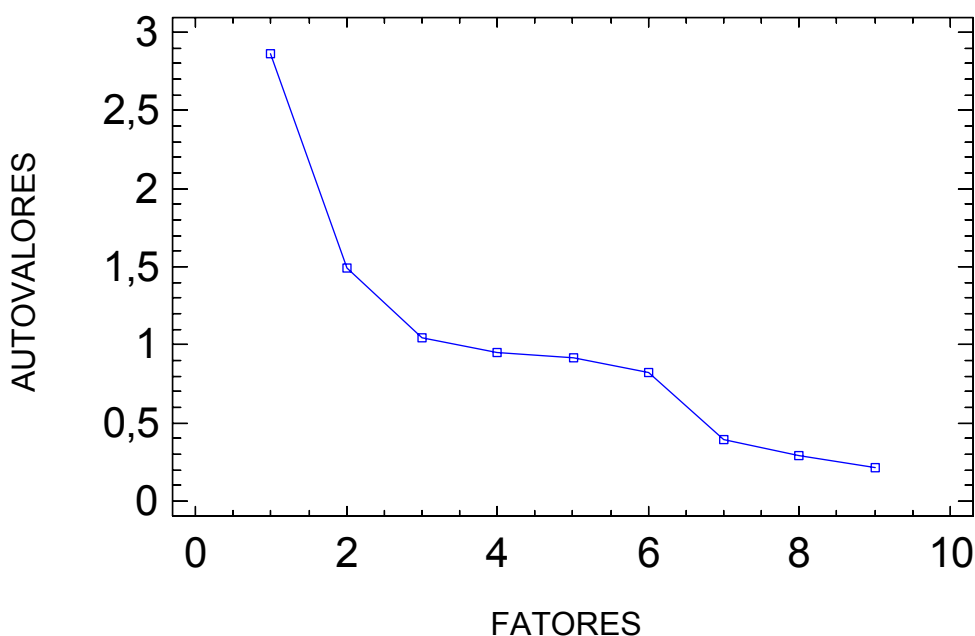


FIGURA 3.1 – Fatores X Autovalores da análise realizada

Na Tabela 3.2 encontram-se as cargas fatoriais que indicam o peso de cada um dos 06 fatores selecionados em cada variável. Por exemplo, a variável 01 é explicada, principalmente, pelos fatores 03, (-0,760365), 04, (0,523348) e 06, (0,302183), seguido dos fatores 02, (-0,157343), 05, (-0,152624) e 01, (0,0826842). Esta tabela mostra também, as equações que estimam os fatores comuns antes da performance de rotação. Por exemplo, o primeiro fator comum tem a seguinte equação:

$$Fator_{01} = 0,0826842 * var_{01} + 0,211078 * var_{02} - 0,331162 * var_{03} + 0,850325 * var_{04} - 0,439593 * var_{05} + 0,653192 * var_{06} + 0,378526 * var_{07} + 0,890015 * var_{08} + 0,652643 * var_{09}$$

variável	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6
01	0,0826842	-0,157343	-0,760365	0,523348	-0,152624	0,302183
02	0,211078	0,143531	0,481553	0,787618	-0,152879	-0,239631
03	-0,331162	0,801976	-0,220391	-0,0694386	0,0496773	-0,0732702
04	0,850325	0,177271	-0,120497	0,0487104	0,0235928	-0,181346
05	-0,439593	0,757301	-0,156562	0,0042705	0,0131147	-0,041758
06	0,653192	0,159154	-0,040330	-0,116783	-0,65197	-0,179519
07	0,378526	0,339168	0,379936	0,042663	-0,0422709	0,767718
08	0,890015	0,219914	0,049093	-0,00674974	-0,0076829	-0,0796584
09	0,652643	0,0965103	-0,055156	0,193948	0,660706	-0,101784

TABELA 3.2 – Matriz Fatorial (antes da rotação)

A soma dos quadrados das cargas fatoriais para cada variável resulta num valor estimado da comunalidade, que é parte da variância total explicada pelos 6 fatores, como pode ser visto na Tabela 3.3.

Variáveis	Comunalidades
1	0,998250
2	0,998185
3	0,814063
4	0,804813
5	0,793192
6	0,924547
7	0,995666
8	0,849349
9	0,922807

TABELA 3.3 – Comunalidades

Assim sendo, nenhuma das variáveis pode ser desconsiderada, todas têm altas comunalidades, ou seja, se aproximam de um, significando que são muito importantes na análise fatorial conduzida; logo, todas as 9 variáveis farão parte tanto no treinamento, como na validação e teste da rede.

### 3.4 – METODOLOGIA PARA RESOLUÇÃO DO PROBLEMA

HAYKIN (2001) apresenta, em termos particularmente simples e claros, as fases necessárias à implantação de uma rede neural:

- a) organizar uma base de dados com toda a informação relevante disponível;
- b) separar esta informação em dois grupos, destinando-se o primeiro à fase de “aprendizagem” da rede, enquanto o segundo será utilizado na fase de teste;
- c) transformar a informação primária referida em “a” de forma a poder servir de entrada à rede;
- d) selecionar, treinar e testar a rede: esta fase, bem como a eventual reformulação das anteriores, deverá ser repetida até que os resultados sejam considerados adequados;
- e) aplicar à rede que se acabou de desenvolver a realidade econômica.

Nesta fase do trabalho foram utilizadas informações disponíveis do cadastro dos associados/clientes, conforme Quadro 3.5, contendo os seus respectivos históricos de empréstimos anteriores, e informações dos atrasos, levando em conta o tempo de atraso, sendo verificado o efeito nas decisões tomadas. Para efeito de análise, são consideradas inadimplentes as operações vencidas e não pagas após o 15º (décimo quinto) dia do vencimento [SICREDI, 2003]. Em se tratando de crédito rural, de acordo com o Decreto-Lei 167/67, considera-se vencida a operação onde conste garantidor inadimplente.

Os candidatos ao crédito foram enquadrados em dois grupos principais: candidatos que foram aceitos, e foram bons, ou seja, honraram com seus pagamentos (adimplentes) e candidatos que foram aceitos, mas faltaram com suas obrigações em seus empréstimos (inadimplentes).

Para treinar adequadamente a rede, a amostra de treinamento deve arcar com uma boa representação da população em estudo [MALHOTRA e MALHOTRA, 2003]. Modelos de redes neurais são bons para generalizar, mas podem ser deficientes na extrapolação. Assim, os dados de treinamento devem abranger a expectativa inteira da entrada do espaço de dados; logo, de acordo com estas diretrizes, a rede foi treinada com uma amostra de 577 observações (vetores). O treinamento colocou uma amostra base com 387 vetores classificados como adimplentes e 190 vetores da classe de inadimplentes, conforme Quadro 3.5 (vide página 47). Além disso, para assegurar que a rede não é treinada somente com vetores de uma classe ou um tipo de crédito, os conjuntos foram estratificados, ou seja, tomou-se o cuidado de se manter a proporção dos 2 conjuntos (adimplentes e inadimplentes).

As arquiteturas das redes foram estabelecidas obedecendo às classes de entrada e saída. Desta maneira, a camada de entrada ficou fixa em 9 (nove) nós de fonte, representando, cada uma, as características dos cooperados, conforme o Quadro 3.4 (vide página 45) e a camada de saída com 1 (um) nó destino representando o fornecimento (1) ou não (0) do crédito.

Finalmente para a rede ser treinada eficientemente, uma amostra com exemplos de cooperados adimplentes, inadimplentes e um número suficiente de neurônios ocultos são também essenciais. Usualmente, o número de neurônios ocultos usado é uma fração relativamente pequena do tamanho da camada de entrada. Muitas regras podem ser aplicadas para solucionar o tamanho da camada oculta [MALHOTRA e MALHOTRA, 2003], mas em muitos casos triviais, o cálculo do erro é a melhor guia. Se a rede falha para convergir para a solução, então mais neurônios ocultos podem ser requisitados. Segundo MALHOTRA e MALHOTRA (2003), uma regra para calcular o número de neurônios ocultos, seria determinado pela equação (3.4.1).

$$\text{número de neurônios oculto} = \left( \frac{\sum \text{entradas} + \sum \text{saídas}}{2} \right) \quad (3.4.1)$$

Uma rede se pouco treinada, resulta em inadequada, mas se muito treinada, pode acontecer da rede se especializar demais nas informações contidas no conjunto de treinamento, perdendo sua capacidade de generalização. Assim, além do número de neurônios ocultos, também foi experimentado com diferentes números de ciclos de treinamento para projetos do modelo da rede neural ótima.

Uma das formas de avaliação do desempenho da rede neural é através do erro médio quadrático (*mse – mean squared error*), isto é, a média do somatório dos quadrados dos erros de cada caso, tanto do conjunto de treinamento, quanto do conjunto de teste. É normal se adotar a raiz quadrada do erro médio quadrático (*rmse*) no cálculo do erro da rede neural. O cálculo do *rmse* pode ser determinado pela equação (3.4.2).

$$rmse = \sqrt{\left( \frac{\sum (y - \tilde{y})^2}{n} \right)} \quad (3.4.2)$$

Onde:

- $y$  = saídas reais normalizadas;
- $\tilde{y}$  = saídas normalizadas previstas pela rede;
- $n$  = número de casos.

Neste trabalho, a implementação das RN's foi realizada com auxílio do programa computacional Matlab 6.5, onde o treinamento das redes foi feito através do algoritmo de Levenberg-Marquadt (LM), uma otimização do algoritmo de *back-propagation*, sendo, em geral, mais eficiente que este último.

### 3.4.1 – VALIDAÇÃO CRUZADA

Na aprendizagem por retropropagação, começa-se com uma amostra de treinamento e usa-se o algoritmo de LM para se determinar os “melhores” pesos sinápticos de um perceptron de múltiplas camadas (*MLP*) alimentando a rede com os 577 exemplos de treinamento. Espera-se que a rede neural projetada se torne

bem treinada de modo que aprenda o suficiente sobre o problema para generalizar sobre ele. BISHOP (1995) destaca que o processo de aprendizagem (treinamento da rede neural) pode ser visto como um problema de “ajuste de curva”.

Em particular, segundo BISHOP (1995), é possível que a rede acabe sendo excessivamente ajustada aos dados de treinamento, se a sessão de treinamento não for encerrada no “ponto” certo. Pode-se identificar o início do excesso de treinamento através do uso da validação cruzada<sup>7</sup>, conforme Figura 3.2.

Segundo STONE, *apud* BISHOP (1995), uma ferramenta padrão da estatística conhecida como validação cruzada fornece um princípio orientador atraente na identificação do excesso da aprendizagem. Primeiramente, o conjunto de dados disponível foi dividido aleatoriamente em um conjunto de treinamento e em um conjunto de teste. O conjunto de treinamento foi dividido em dois subconjuntos disjuntos, conforme Quadro 3.5 (vide página 47).

- *Subconjunto de estimação*, usado para selecionar o modelo;
- *Subconjunto de validação*, usado para validar o modelo.

A motivação aqui é validar o modelo com um conjunto de dados diferente daquele usado para estimar os parâmetros. Há, entretanto, uma possibilidade considerável de que o modelo assim selecionado, com os valores de parâmetros com melhor desempenho, possa acabar ajustando excessivamente o subconjunto de validação.

Para resguardar esta possibilidade, o desempenho de generalização do modelo selecionado foi medido sobre o conjunto de teste, que é diferente do conjunto de validação.

A Figura 3.2 mostra formas conceituais de duas curvas de aprendizagem, uma relativa às medidas sobre o conjunto de estimação (treinamento) e a outra

---

7 – A história do desenvolvimento da validação cruzada está documentada em Stone (1974).

relativa ao subconjunto de validação (teste). Tipicamente, o modelo não funciona tão bem sobre o subconjunto de validação do que sobre o conjunto de estimação, sobre o qual o projeto foi baseado. A curva de aprendizagem de estimação decresce monotonamente para um número crescente de épocas, da maneira usual. Diferentemente, a curva de aprendizagem de validação decresce monotonamente para um mínimo e então começa a crescer conforme o treinamento continua.

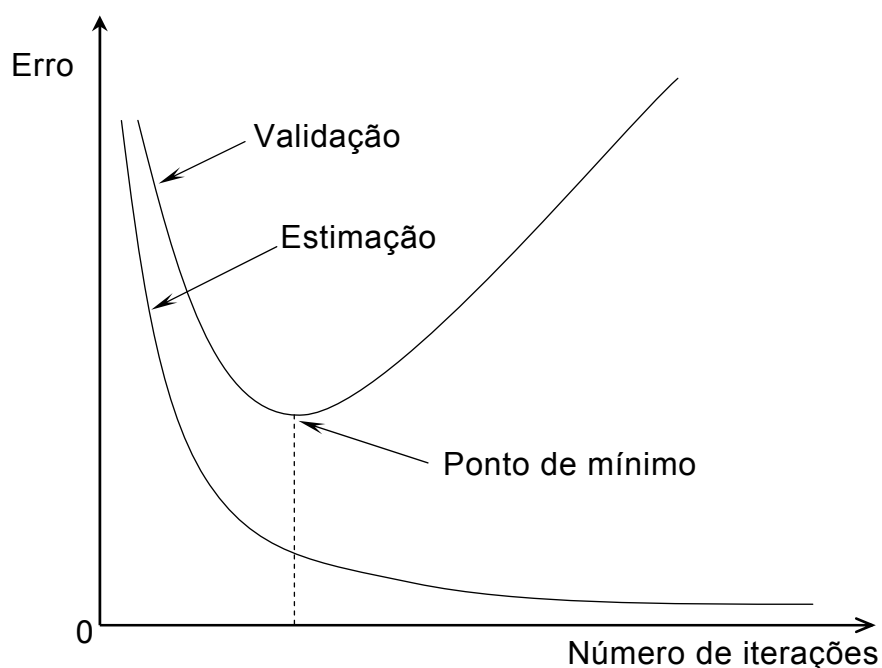


FIGURA 3.2 – Treinamento versus capacidade de generalização da RN

FONTE: HAYKIN, 2001.

Ao analisar-se a curva de aprendizagem de estimação pode parecer que poder-se-ia melhorar o desempenho indo além do ponto mínimo da curva de aprendizagem de validação. Na realidade, entretanto, o que a rede aprende após este ponto, segundo BISHOP (1995), é essencialmente o ruído contido nos dados de treinamento. Esta heurística sugere que o ponto mínimo na curva de aprendizagem de validação seja usado como critério sensato para encerrar a sessão de treinamento.

### 3.5 – MODELOS DESENVOLVIDOS

Muito embora vários modelos possam ser desenvolvidos com vista à tomada de decisões em um ambiente de complexidade crescente, neste estudo é utilizada a abordagem proposta por MALHOTRA e MALHOTRA (2003): a complexidade inerente à interação entre os dados estruturais de um sistema, associada à possibilidade de erro inerente à incerteza decorrente de uma análise específica, deve ser tratada como um modelo de “caixa preta”, ou seja, assume-se que não seria de fato plausível que uma análise qualquer fosse a mais adequada (o que de fato pode ser um argumento em favor da personalização). Entretanto, alternativas das mais diversas teriam sua efetividade observada pelos resultados obtidos independentemente da metodologia empregada. Aqui, nota-se a possibilidade alternativa oferecida pelos modelos de redes neurais artificiais.

Dentre as alternativas relativas especificamente ao método relacionado às diversas abordagens sistêmicas de construção dos referidos cenários (como por exemplo, o emprego da Análise Multivariada), os modelos de redes neurais artificiais destacam-se como uma abordagem numérica e de processamento paralelo que apresentam resultados bastante positivos na busca e classificação de padrões de dados semelhantes, e que desde os anos 80 tem ganhado crescente popularidade entre o meio acadêmico e o mundo dos negócios.

O Teorema da Aproximação Universal [HAYKIN, 2001] pode ser visto como um teorema existencial que fornece a justificativa matemática para a aproximação de uma função contínua arbitrária, em oposição à representação exata. Este teorema afirma que uma única camada oculta é suficiente para um perceptron de múltiplas camadas (*MLP*) computar uma aproximação uniforme para um dado conjunto de treinamento, representado pelo conjunto de entradas e a saída desejada (alvo). O interesse na pesquisa sobre as “virtudes” dos *MLP* como dispositivo para a representação de funções contínuas arbitrárias é invocada de uma versão melhorada do Teorema da Superposição de Kolmogorov [HAYKIN, 2001]; decidiu-se no projeto desenvolvido pela utilização de somente uma camada oculta, com o

número de neurônios variando de 2 a 11; e a camada de saída foi composta por apenas um neurônio, onde, as redes foram projetadas para fornecer como saída o *status* do cooperado, ou seja, adimplente ou inadimplente.

Em termos básicos, a diferenciabilidade é a única exigência que a função de ativação deve satisfazer, pois a função de ativação empregada deve necessariamente ser contínua. Neste estudo, na camada oculta foi empregado como função de ativação não-linear, continuamente diferenciável, as funções tipo sigmoidal, pois uma função sigmóide assume um intervalo contínuo de valores entre 0 e 1 e ela é definida como uma função estritamente crescente que exibe um balanceamento adequado entre comportamento linear e não-linear. Portanto, nos neurônios da camada oculta, foram adotadas como função de ativação sigmoidal, *logsig* ou *tansig*. Já a camada de saída, composta de um neurônio, teve como função de ativação linear (*purelin*) ou sigmoidal (*logsig*, *tansig*).

O algoritmo de treinamento utilizado foi o de Levenberg Marquardt, sendo as redes projetadas para calcular os pesos sinápticos do perceptron de múltiplas camadas, isto é, aprendizagem por retropropagação do erro. Segundo HAYKIN, 2001, um perceptron de múltiplas camadas treinado com o algoritmo de retropropagação pode ser visto como um veículo prático para realizar um mapeamento não-linear de entrada-saída de natureza geral.

O estudo começa com dois neurônios na camada oculta como ponto de partida para os resultados de simulação resumidos na Tabela 4.1 (vide página 59), sendo que foi adotado um número máximo de 100 épocas para o treinamento. Para assegurar que a rede não é treinada somente com vetores de uma classe ou um tipo de crédito, foi incluído, aleatoriamente, um número de vetores pertencentes às classes de adimplência e inadimplência. Sendo a ordem de apresentação aleatória de época para época.

## CAPÍTULO IV

### 4.0 – OBTENÇÃO DOS RESULTADOS

Neste capítulo é efetuada uma análise da viabilidade da aplicação de um método computacional totalmente baseado em modelos de redes neurais de aprendizado supervisionado, através da utilização de uma base de dados relativa ao Sistema de Crédito Cooperativo. O método utilizado baseia-se em um processamento efetuado em fases, feito seqüencialmente. Cada uma das fases se torna geradora de entradas para a fase seguinte; contudo, o tempo necessário no desempenho de cada uma das tarefas é relativamente baixo, especialmente quando comparado às metodologias mais tradicionais.

A nomenclatura adotada neste estudo, que será utilizada para representar as redes, tem as seguintes características, pela ordem: número de parâmetros de entrada, que ficou fixa em 9 nós de fonte; função de ativação do neurônio da camada oculta; número de neurônios na camada oculta; número de iterações do treinamento; função de ativação do neurônio da camada de saída. Assim sendo, neste formato a denominação das redes começa com um número correspondente à quantidade de variáveis de entrada, seguido das letras “Et” ou “El” representando a função de ativação *tansig* e *logsig*, respectivamente, para os neurônios da camada oculta; em seguida, tem-se o número de neurônios da camada oculta seguido do número de iterações (máximo 100) e seguido, finalmente, das letras “p” ou “l” ou “t” representando a função de ativação *purelin*, *logsig* e *tansig* respectivamente.

As redes dispõem de uma camada oculta, onde um número representa o número de neurônios desta camada seguido da letra “n”. Por exemplo, a rede 9Et7n61p é uma rede com 9 nós de fonte, com a camada oculta composta de 7 neurônios, tendo como função de ativação *tansig*, sendo treinada com 61 épocas, tendo como função *purelin* no neurônio da camada de saída e foi treinada com o

algoritmo “*trainlm*”. A entrada na rede foi aleatória, de época para época, e os valores normalizados. A Tabela 4.1 resume as melhores redes conforme as configurações acima especificadas.

Conforme já mencionado, uma maneira informativa de examinar o desempenho das redes é através da raiz quadrada do erro médio quadrático (*rmse*) de treinamento, de teste e pela frequência de casos com menores erros. As redes foram treinadas com os dados de entrada e saída normalizados para a faixa de intervalo [0 1] e o resultado do erro médio quadrático (*mse*), após a extração da raiz quadrada, foi multiplicado por 100, para se obter o valor percentual. O *rmse*, desta forma, representa a porcentagem de erro em relação à classificação correta dos cooperados. Os *rmse*’s foram determinados durante o processo através do programa computacional Matlab 6.5, calculados pela Equação (3.4.2).

A Tabela 4.1 apresenta as redes consideradas de melhor desempenho por mostrarem os menores percentuais de erros, tanto de teste como de treinamento, sendo levado em consideração o número de épocas conforme a teoria de validação cruzada. Nas 7 redes da Tabela 4.1, os *rmse*’s de treinamento se situaram no intervalo entre 2,47% e 9,32%, e os *rmse*’s de teste ficaram no intervalo entre 4,86% e 12,13%. Estes resultados indicam que estas redes tiveram um desempenho satisfatório no número de iterações para treinamento.

TABELA 4.1 Percentuais de erros obtidos nas simulações para um nº de variável de neurônios ocultos

REDE	Erro médio quadrático <i>rmse</i> (%)		Média entre OS ERROS
	TREINAMENTO	TESTE	
9Et5n93p	4,10	10,44	7,27
9El6n88p	2,47	12,13	7,30
9Et7n63l	3,26	9,46	6,36

Continua

			Continuação
9Et8n70t	4,20	5,12	4,66
9EI9n59I	9,32	8,54	8,93
9Et10n65I	6,22	4,86	5,54
9EI11n48I	5,89	8,55	7,22

TABELA 4.1 Percentuais de erros obtidos nas simulações para um nº de variável de neurônios ocultos

Examinando as redes desenvolvidas, observa-se que um erro médio quadrado pequeno não implica necessariamente em boa generalização (isto é, bom desempenho com dados não-vistos anteriormente), pois as redes que apresentaram *rmse* de treinamento abaixo dos valores apresentados para as redes da Tabela 4.1 tiveram alto *rmse* de teste, e vice-versa, o que prejudicou seu desempenho. Este fenômeno de baixo valor de *rmse* de treinamento e alto valor de *rmse* de teste é conhecido como excesso de ajuste ou excesso de treinamento; quando a rede é treinada em excesso, ela perde a habilidade de generalizar entre padrões de entrada-saída similares. Com o intuito de especificar corretamente o treinamento, ou o mais próximo possível, foi introduzido uma ferramenta padrão de estatística conhecida como Método de Validação Cruzada, conforme já mencionado. Na prática esse método tende a identificar o início do excesso de treinamento, conforme destacado anteriormente, em treinamento versus capacidade de generalização da rede neural, Figura 3.2 (vide página 55).

### 4.3 – DESEMPENHO DA MLP

Refletindo a abordagem prática para o problema da determinação do número ótimo de neurônios ocultos, o critério utilizado foi o menor número de neurônios ocultos que produz um melhor desempenho para rede, isto é, melhor média referente aos erros de treinamento e teste, da respectiva rede.

Como os pesos sinápticos da rede devem ser ajustados para minimizar o erro médio quadrático entre a saída-alvo (resposta desejada) e a saída real da rede, foi introduzido para uma melhor performance, uma função com o intuito de melhor aproximar a saída real da rede da saída-alvo, lembrando que a saída-alvo é composta de zero ou um (adimplente ou inadimplente). Este artifício necessário, implementado através do programa computacional Matlab, deve-se ao fato de redes que apresentam o neurônio da camada de saída com função de ativação *purelin* podem, eventualmente, produzir resultados diferentes do intervalo [0 1].

A Figura 4.1 ilustra o melhor valor de corte para a rede 9EI6n88p, que aproximadamente ficou em 0,65, ou seja, valores menores ou iguais a 0,65 são automaticamente reconhecidos como zero e valores superiores a 0,65 são reconhecidos como um.

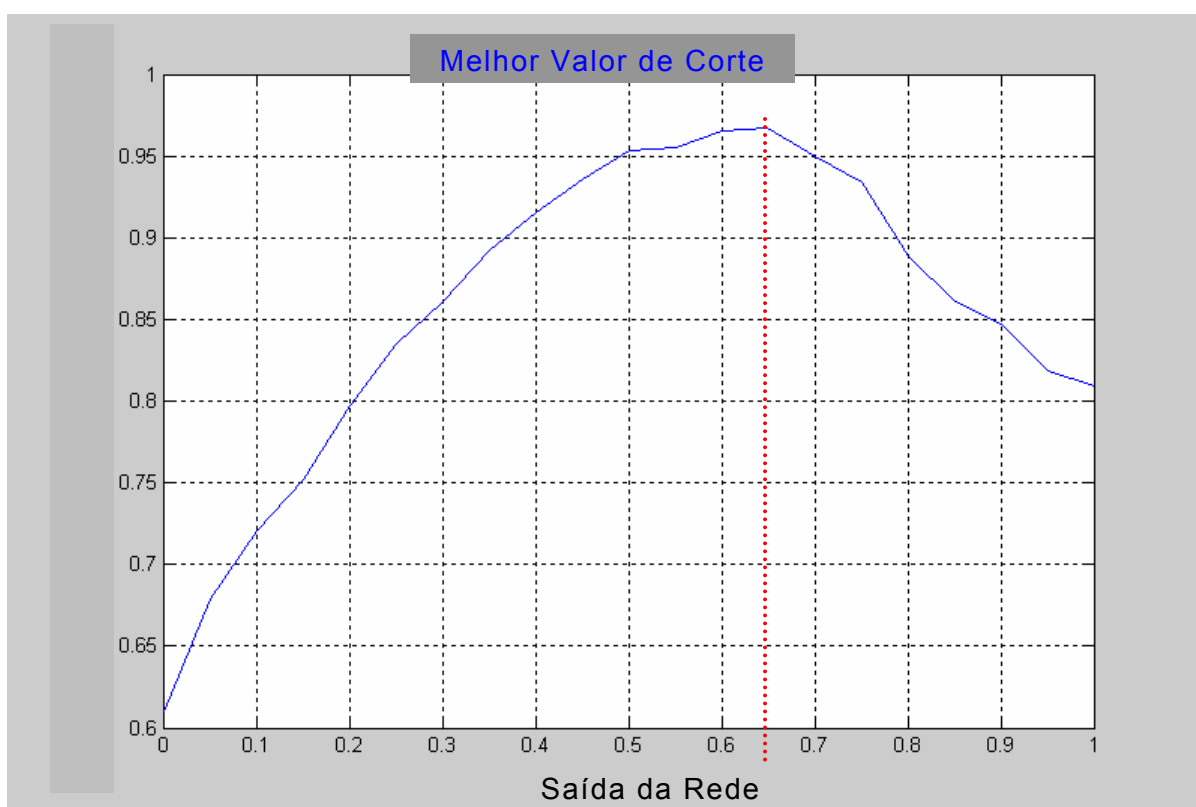


FIGURA 4.1 – Ilustração para a identificação do melhor valor de corte, neste caso para a rede 9EI6n88p.

O desempenho dos modelos das RN's também pode ser observado graficamente, tanto para o conjunto de treinamento como para o conjunto de teste. A figura 4.2 mostra a rede 9Et8n70t na fase de treinamento, com 70 épocas, tendo como função de ativação *tansig* tanto na camada oculta quanto na camada de saída, sendo disposta de 8 neurônios ocultos. Esta rede teve uma das melhores performances entre as redes desenvolvidas.

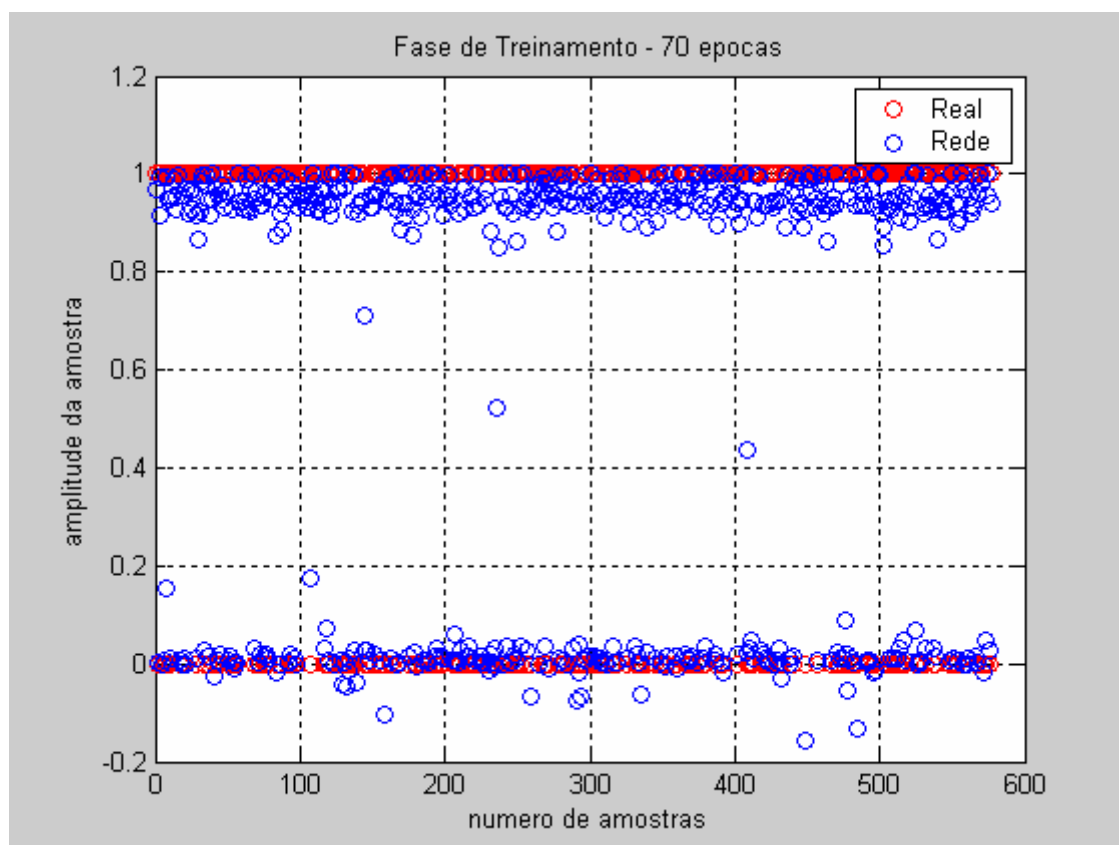


FIGURA 4.2 – Ilustração da performance da rede na fase de treinamento. Neste caso, da rede 9Et8n70t.

Os gráficos relacionam o número real de cooperados adimplentes com o número previsto, correspondentes aos conjuntos de treinamento e teste respectivamente. Estes gráficos apresentam no eixo **x** (horizontal) o número total de cooperados apresentados à rede (tanto inadimplentes como adimplentes) e no eixo **y** (vertical) a amplitude das amostras normalizadas no intervalo [0 1]. O neurônio da camada de saída tem o propósito de aplicar uma seqüência de ajustes corretivos

aos pesos sinápticos. Estes ajustes tendem a aproximar o sinal de saída da resposta desejada.

Estes gráficos, como o da Figura 4.3, têm a finalidade de melhor definir e classificar as garantias necessárias à classificação do melhor projeto de rede. Neste ponto já é possível visualizar todo o conjunto de dados que a rede utilizou, e chegar à conclusão: se ela está correta, se não falta ainda uma calibragem nos dados, se alguma preferência (peso) da avaliação geral do cooperado não precisa ser revista, e se a estatística não está desatualizada.

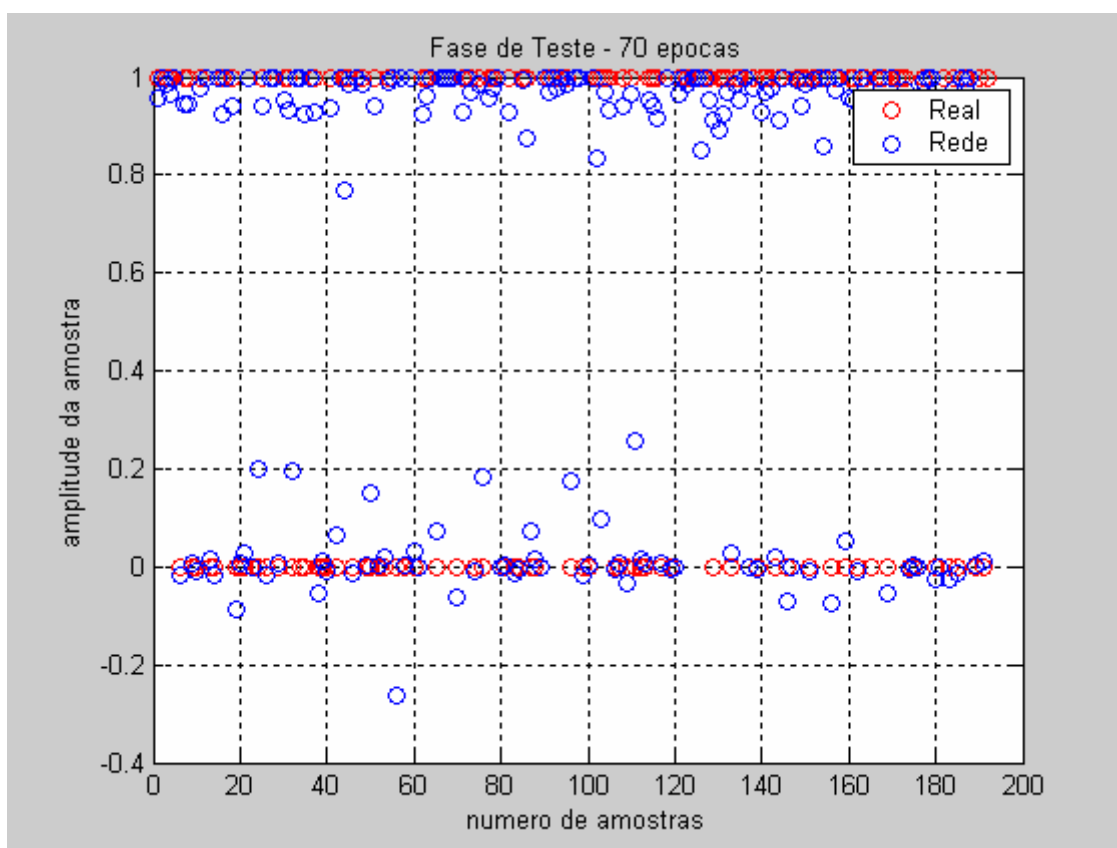


FIGURA 4.3 – Ilustração da performance da rede na fase de teste. Neste caso, da rede 9Et8n70t.

Em complementação à análise do desempenho das redes, conforme mostram os gráficos agrupados nas Figuras 4.4 e 4.5, cada analista pode chegar a conclusões diversas, analisando cada rede individualmente. Deve-se tomar um cuidado extra nestas análises isoladas levando-se em consideração que os valores foram colocados de acordo com a classificação de cada cooperado, conforme

Quadro 3.4 e posteriormente, antes da primeira apresentação à primeira rede, normalizados no intervalo de [0 1]; logo, estando todos num mesmo patamar, a análise através do erro ainda se destaca como sendo a melhor guia.

Dentre as redes constantes na tabela 4.1, a rede 9Et8n70t teve desempenho levemente superior quando considerado a frequência de casos com erro de teste inferior a 6%; já a rede 9Et7n63l teve um desempenho bom se considerado somente o valor de treinamento (3,26%), porém o erro na fase de teste ficou muito a desejar 9,46%.

A visualização dos resultados proporcionada pelas Figuras 4.2 à Figura 4.5 se mostrou interessante do ponto de vista da disponibilidade da ilustração de suas estruturas, porém a forma de análise dos resultados gera uma tabela de coincidência entre os dois padrões, o que caracteriza uma confusão de resposta, o que torna seu exame uma tarefa mais objetiva, reforçando que a análise através do erro é o melhor caminho.

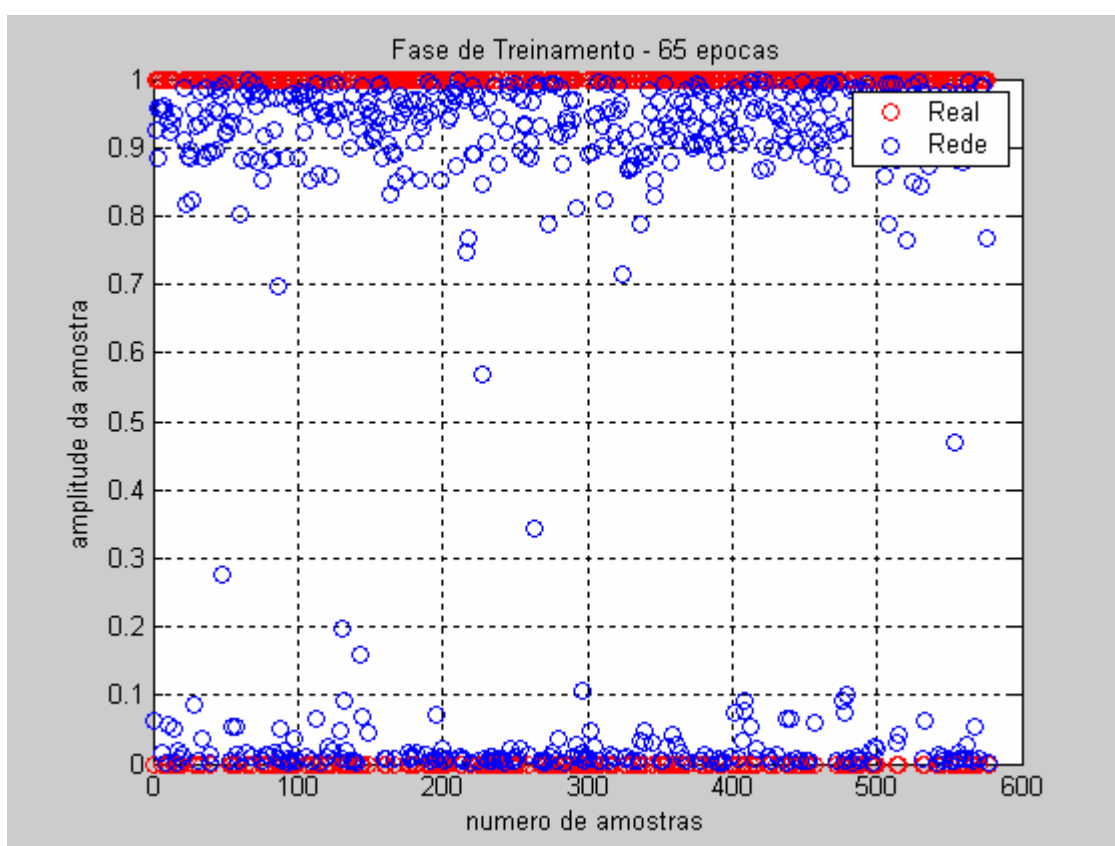


FIGURA 4.4 – Ilustração da performance da rede na fase de treinamento. Neste caso, da rede 9Et10n65l.

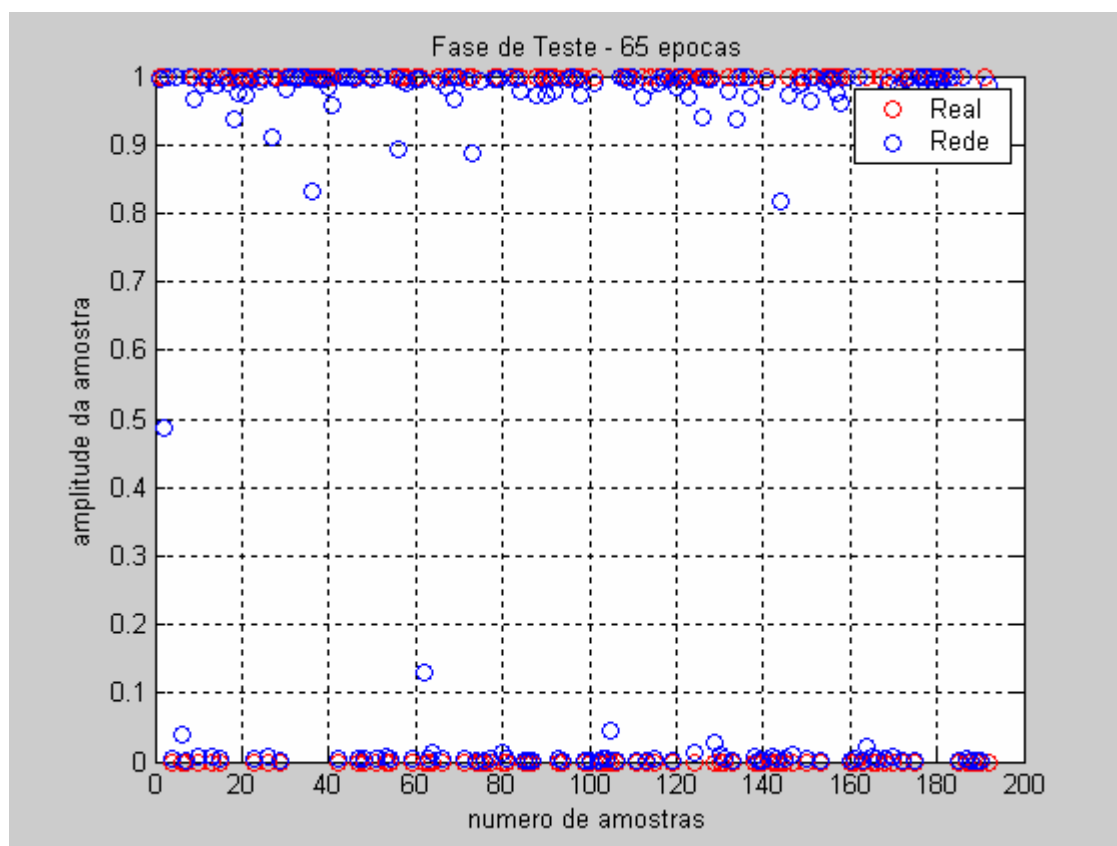


FIGURA 4.5 – Ilustração da performance da rede na fase de teste. Neste caso, da rede 9Et10n65l.

Observa-se, também, que a classificação como satisfatória das redes é correta para o conjunto de teste, já que na análise do desempenho da rede os resultados vêm para corroborar com a avaliação de uma melhor classificação.

#### 4.4 – COMPARAÇÃO DE DESEMPENHO DOS MODELOS

De uma forma sucinta, a análise dos resultados centrou-se em identificar a capacidade das redes neurais de apresentar como saída um indicativo de adimplente ou inadimplente, dado um conjunto de informações de entrada. Os treinamentos obedeceram a uma criteriosa distribuição na escolha de amostras, tendo em vista o estabelecimento de conjuntos de dados.

As variáveis utilizadas no treinamento, validação e teste das redes passaram por um pré-estudo de caso para uma melhor definição na representação do cooperado escolhido. Nas redes destacadas com o melhor desempenho foi aplicado o teste final, que consiste em avaliar os cooperados que obtiveram crédito e, por causas não estudadas neste trabalho, não honraram com os devidos compromissos assumidos, constando agora como inadimplentes. Deve-se Lembrar que os candidatos ao crédito foram caracterizados em dois grupos principais: candidatos que foram aceitos e foram bons, isto é, honraram com seus pagamentos (adimplentes) e candidatos que foram aceitos, mas faltaram com suas obrigações em seus empréstimos (inadimplentes).

Basicamente a avaliação final do desempenho das RN's consistiu em localizar na base de dados, utilizadas para treinar e testar as redes, os cooperados que no momento da aprendizagem e teste constavam como adimplentes, isto é, foi fornecido crédito ao cooperado. Porém, exaurido o prazo para pagamento se posicionaram como inadimplentes. Localizados estes cooperados e após a rede ter sido calibrada é hora de efetuar a avaliação final que consiste em repassar nestas redes, os cooperados tendo agora como saída da rede zero (não fornece o crédito) e não mais o valor um, como na fase de treinamento e teste; a análise do desempenho da rede foi efetuada através do erro médio quadrático entre a saída-alvo (resposta desejada) e a saída real da rede, na classificação destes cooperados.

Foi assumido para o número de inadimplentes, conforme descrito no parágrafo anterior, uma frequência percentual que se encontra na faixa de 5% até 13% do total de cooperados analisados. Para a rede 9Et8n70t, conforme Figura 4.6, foi obtido em média um acerto de 18,49%, ou seja, de cada 100 cooperados a classificação mostrou-se correta em mais de 18 cooperados se comparado ao sistema corrente, que é o atualmente utilizado no Sistema de Crédito Cooperativo. É verdade que o desempenho do modelo aqui desenvolvido foi superior ao sistema corrente, pela melhor classificação disponível até no visual, conforme Figuras 4.7 e 4.8, porém, a superioridade não é tão significativa.

De forma a melhor verificar os resultados da classificação, obtidos através do erro médio quadrático, a Figura 4.6 apresenta uma comparação, de forma ajustada, os resultados obtidos.

### *A D I M P L Ê N C I A*

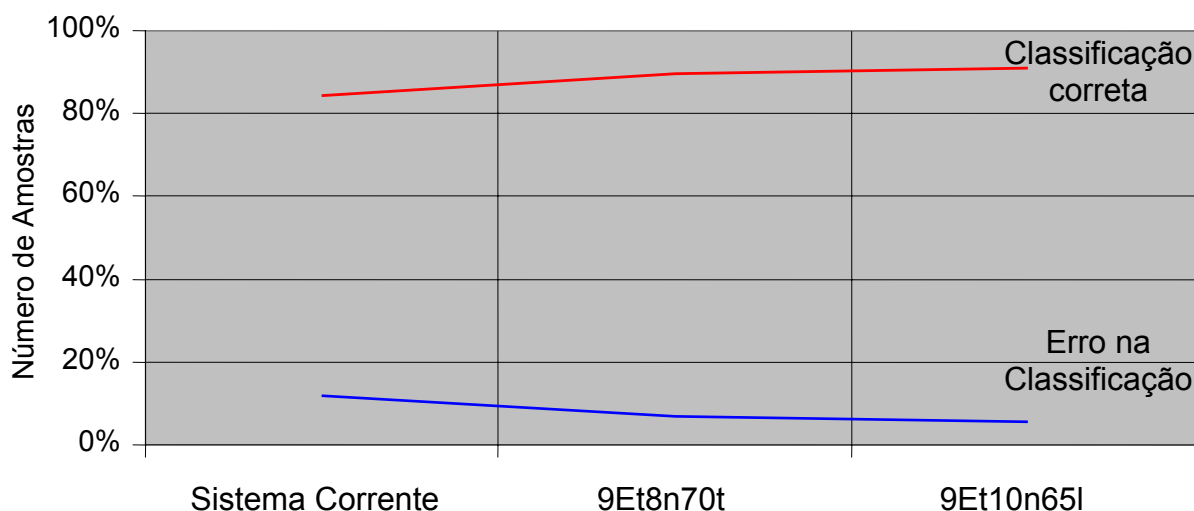


FIGURA 4.6 – Acertos das classes de saída por arquitetura.

Como pode ser observado na Figura 4.6, há um comparativo entre o desempenho de classificação do sistema corrente com as redes 9Et8n70t e 9Et10n65l, mostrando em cada uma a classe de resultados. Nesta apresentação foram utilizados os valores do sistema corrente e das redes, que obtiveram melhor desempenho, ou seja, 87,20%, 89,60% e 90,22% respectivamente, referentes a correta classificação. E, conseqüentemente, tem-se como erro na classificação dos cooperados os percentuais 12,80%, 10,40% e 9,78% relativos, respectivamente, ao sistema corrente e as redes 9Et8n70t e 9Et10n65l.

A rede 9Et10n65l, conforme Figura 4.7, na previsão de inadimplência aos cooperados mostrou-se superior tanto ao sistema corrente, tendo uma média de aproximadamente 3,02% de acertos, além; quanto à rede 9Et8n70t, que registrou uma melhora na performance de 0,62%.

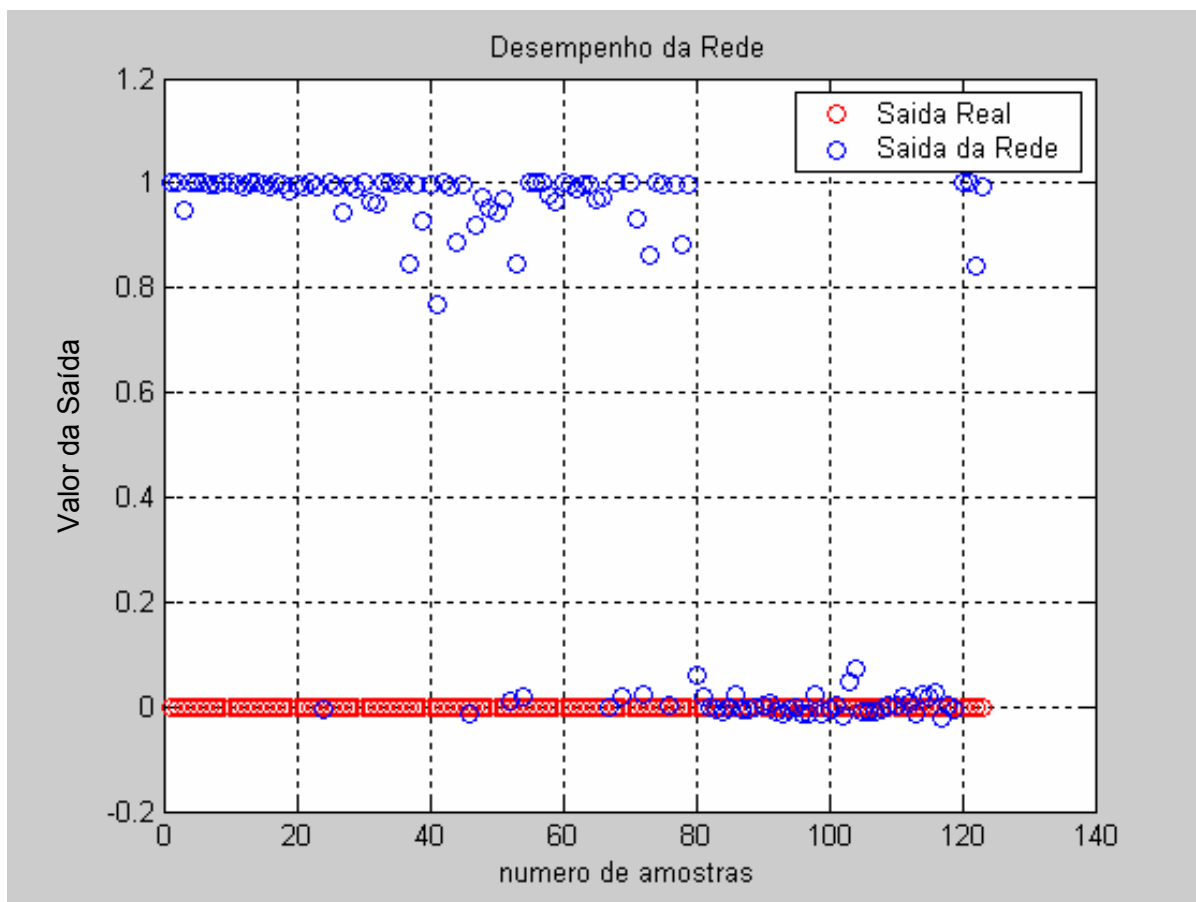


FIGURA 4.7 – Ilustração da performance no desempenho da rede. Neste caso, da rede 9Et10n65l.

Evidentemente, é preferível uma maior frequência de casos em que o erro na classificação é menor e, uma menor frequência de casos em que esse erro é maior.

A ilustração da performance no desempenho da rede 9Et8n70t pode ser visualizado através da Figura 4.8, que apresentou uma discreta melhora na previsão de inadimplência aos cooperados se comparado ao sistema corrente, tendo um percentual de aproximadamente 2,40% de acertos, além, dos classificados corretamente pelo sistema atual.

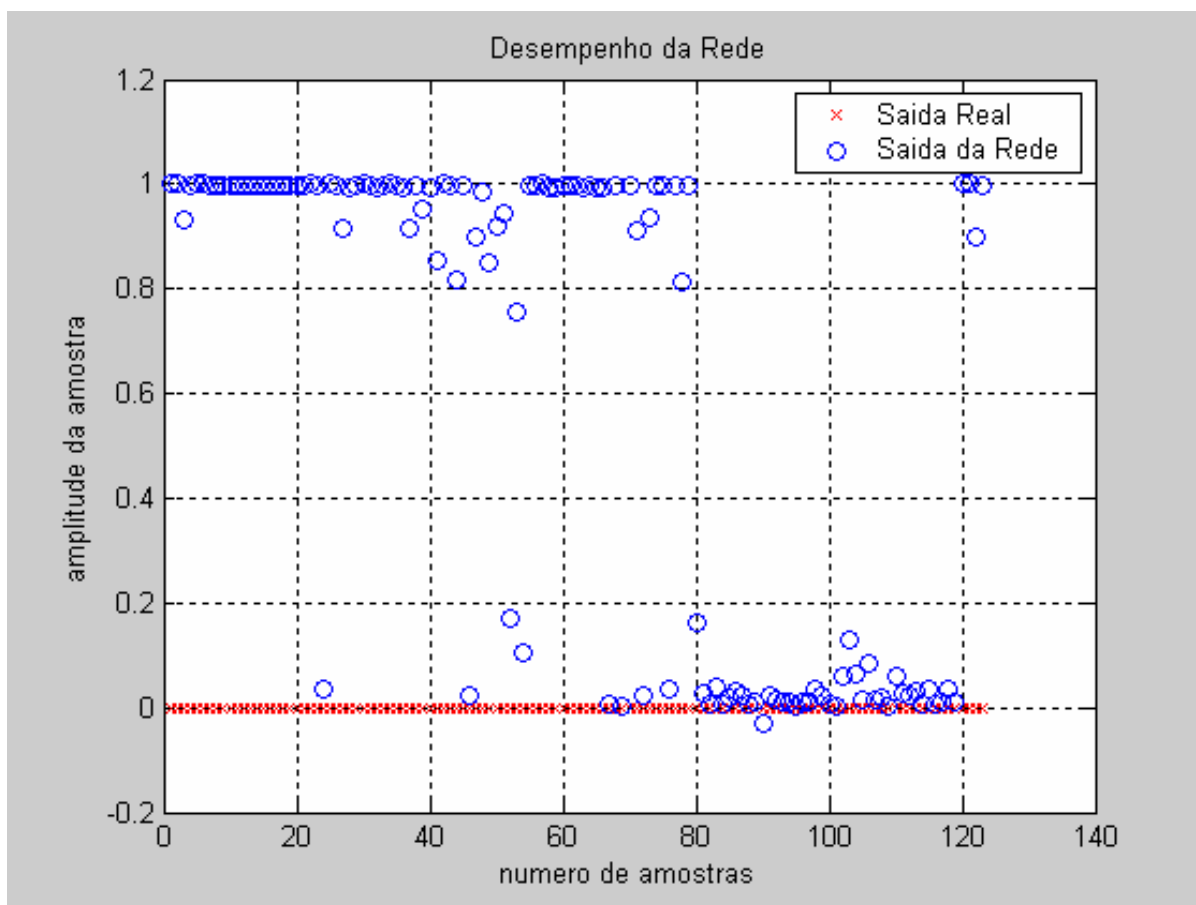


FIGURA 4.8 – Ilustração da performance no desempenho da rede. Neste caso, da rede 9Et8n70t.

A provisão para devedores inadimplentes é outra fonte de encarecimento do crédito e é claro que a inadimplência não é um evento binário do tipo “dívida paga” ou “devedor inadimplente”. Em primeiro lugar, a cada espécie de contrato vincula-se uma espécie de garantia, ou mais de uma. Além disso, não se deve confundir a inadimplência com a simples mora. A renegociação de dívidas é prática corriqueira. Não raro o pagamento é feito, embora não no tempo correto nem no montante total, mas representando uma perda pouco significativa, especialmente se as condições da renegociação envolverem apenas alongamento de prazo e desconto sobre multas e juros moratórios.

Seguramente o ponto de discussão mais intensa acerca dos resultados obtidos gira em torno da própria adoção de uma metodologia de processamento. Isso também pode ser apreciado sob uma ótica favorável na qual o mais importante

atributo da realização de uma tarefa qualquer seria a obtenção do melhor desempenho possível, independentemente de qual a metodologia adotada.

Observa-se, na prática, com referência a uma vastíssima bibliografia na área computacional, que a utilização de modelos neurais no controle e monitoramento de sistemas complexos geralmente apresenta resultados mais positivos do que aqueles obtidos por técnicas matemáticas tradicionais. Tanto em relação a uma maior eficácia na realização das tarefas, quanto no menor dispêndio computacional e na necessidade de grande especialização do operador, redundando em uma crescente utilização desta metodologia em problemas computacionais e de engenharia.

## CAPÍTULO V

### 5.1 - CONCLUSÕES

Neste estudo verificou-se que as redes neurais através do algoritmo de treinamento Levenberg Marquadt com uma camada oculta, foram capazes não somente de reproduzir o modelo atualmente empregado no Sistema de Crédito Cooperativo, mas também demonstrar uma superioridade, conforme é destacado nas Figuras 4.6 e 4.7. O trabalho é relevante por comparar uma classe que tem obtido bons resultados na literatura (os modelos estatísticos), com uma classe de modelos flexíveis (RN's) ainda pouco explorada nos estudos sobre cooperativas de créditos a pessoa física mostrando que esta classe de modelos flexíveis é capaz de obter um desempenho superior na maioria dos casos considerados. Os resultados obtidos corroboram o estudo, divulgando que os modelos de RN em grande parte dos casos reproduzem satisfatoriamente as características do modelo atualmente empregado.

Foi também ponto importante neste estudo o emprego da validação cruzada, que teve como intuito uma aproximação para a obtenção do número de épocas ótima, tornando possível avaliar nos modelos estimados, o ponto de parada do treinamento. Isto serviu para que a rede neural projetada se tornasse bem-treinada, de modo que aprendesse o suficiente sobre o problema para generalizá-lo, pois caso a rede acabasse sendo excessivamente ajustada aos dados de treinamento, o erro percentual para avaliação seria com certeza maior.

É importante destacar que o desenvolvimento das RN's foi limitado face ao conjunto de dados não ter uma dimensão mais expressiva, no sentido de ser possível a inclusão de variáveis tidas como importante na análise de crédito, sendo destacada a falta de análise em um resumo financeiro do cooperado, onde conste

sua situação patrimonial. Porém, mesmo na ausência deste e de outros dados, as redes desenvolvidas tiveram uma performance bastante satisfatória.

Especificamente no caso de dados cooperativos as redes neurais se mostraram ferramentas que proporcionam ao usuário uma comparação intuitiva bastante interessante, notadamente na obtenção de estruturas visuais, conforme Figuras 4.2 a 4.5, que indicam tanto uma categorização relativa aos grupos distintos que compõem cada categoria, quanto o posicionamento dos elementos dentro de um mesmo grupo.

Os modelos de redes neurais apresentam boa performance se comparados a modelos estatísticos. Conforme vasta bibliografia, aquelas podem oferecer um modelo competitivo aproximado para avaliar o cooperado na concessão de empréstimos. Mas há algumas limitações que restringem o uso destas redes na classificação. Uma destas limitações seria a topologia da rede; não há uma teoria formal para determinar a topologia ótima da rede neural para uma aplicação na classificação. Para esse fim, decisões que aproximam, tal como o número de camadas e tamanho da camada oculta precisam ser determinados, utilizando experimentações. Em suma, existe a possibilidade de subestimar ou superestimar a rede, e o resultado é sensível na seleção do parâmetro de aprendizagem. Resultados não precisos podem também ocorrer se a função de transferência selecionada não for adequada. Assim, o desenvolvimento e a interpretação nos modelos de redes neurais requerem mais experiência do usuário do que nos modelos tradicionais estatísticos.

Considerando-se as diferenças entre as características básicas e formas de resolver um problema por um especialista humano e um especialista artificial, chega-se à conclusão de que a aparente limitação de dados manipulados e disponíveis para análise das redes é natural, uma vez que a própria forma de se construir esta classe de redes busca limitar e restringir a abrangência aos fatos essenciais para a solução dos problemas na maioria dos casos e situações conhecidas. Toda vez que surge algum fato novo que afete os resultados e o bom funcionamento das redes, estas devem ser “recalibradas” (sofrer atualização em sua base de conhecimento).

Em suma, apesar deste trabalho apresentar o bom desempenho das redes desenvolvidas, no sentido de terem tido êxito nas análises das classificações dos cooperados, não foi desenvolvido um sistema computacional e, portanto o mesmo deve ser encarado somente como um “protótipo” não sendo ainda suficientemente robusto, completo e funcional para ser utilizado no Sistema de Crédito Cooperativo, que disponibilizou as informações que permitiram desenvolver este estudo.

## 5.2 – SUGESTÕES PARA ESTUDOS FUTUROS

Seguem abaixo algumas sugestões para a continuidade de pesquisa na área, tendo em vista que vários aspectos pertinentes não são abordados pelo presente trabalho.

- Utilização de arquiteturas capazes de realizar processamento de dados difusos (lógica *fuzzy*), o que permitiria a geração de categorias superpostas, e uma posterior comparação dos resultados com a metodologia aqui empregada;
- Desenvolvimento de outros tipos de RN's, como as redes de função de base radial, ou mesmo, o emprego de outros processos de treinamento, adoção de outras funções de ativação dos neurônios, e utilização de técnicas de extração de regras por RN's;
- Uma outra possível análise seria a divisão do erro percentual em faixas, determinando a frequência de casos por faixa de erro tanto para treinamento como para teste, isto, na análise do desempenho das redes por faixas de erro. Este tipo de análise consta em MADALOZZO, 2003.
- Comparar diretamente a utilização de sistemas de classificação neural com as técnicas mais comuns mencionadas, através da apreciação de suas potencialidades e inadequações intrínsecas;
- Outras aplicações de interesse nas áreas relativas à Economia e Finanças, especialmente relacionadas ao estudo de risco e composição de carteiras de ativos (portfólio), tendo em vista a capacidade de identificação de relacionamento a posteriori observada no proposto neste estudo.

Além destas possibilidades, sem dúvida o desenvolvimento de um sistema computacional seria muito útil para efetuar classificações de dados relativos a uma

vasta gama de aplicações, como planejamento urbano e regional, planejamento cooperativo, organização industrial e economia de empresas, análise de concorrência, estudos em Contabilidade Social e Desenvolvimento Econômico, pesquisa e análises sociais e demográficas, e muitas outras.

## REFERÊNCIAS BIBLIOGRÁFICAS

- [01] – ARRAES, D.; SEMOLINI, R.; PICININI, R. **Arquiteturas de Redes Neurais Aplicadas a Data Mining no Mercado Financeiro. Uma Aplicação para a Geração de Credit Ratings** – V Congresso Brasileiro de Redes Neurais. Rio de Janeiro, 2001.
- [02] – ATIYA, A. F. **Bankruptcy Prediction for Credit Risk Using Neural Network: A Survey and New Results**. California Institute of Technology, Pasadena, 2001 IEEE.
- [03] – BACK, J. and CHO, S. **Bankruptcy Prediction for Credit Risk Using an Auto-Associative Neural Network in Korean Firms**. Department of Industrial Engineering, Seoul National University, 2003 IEEE.
- [04] – BANSICREDI: **Manual de Políticas de Crédito** – Edição: Bansicredi – Gerência de Crédito. Abril, 2002.
- [05] – BARRETO, J. M.; AZEVEDO, F. M.; BRASIL, L. M. **Uma Arquitetura Híbrida para Sistemas Especialistas**, GPEB, UFSC, 2003.
- [06] – BISHOP, C. M.; **Neural Networks for Pattern Recognition** - Oxford University Press, 1995.
- [07] – BUSCEMA, M. and SACCO, P. L. **Feedforward networks in financial predictions: the future that modifies the present**. Expert Systems, vol. 17, n. 3, 2000.

- [08] – CARDOSO, R. L. **Estudo exploratório do Processo decisório de Crédito para Pessoas Físicas nas Instituições Financeiras**, Dissertação de Mestrado, São Paulo, FGV / EAESP, 2000.
- [09] – CHATFIELD, C. COLLINS, A. J. **Introduction to Multivariate Analysis**. – School of Mathematics Bath University, Chapman & Hall, London, Reprinted – 1992.
- [10] – DYMINSKI, A. S. **Análise de problemas geotécnicos através de redes neurais**. Tese de Doutorado, PUC – Rio de Janeiro, 2000.
- [11] – FRANCO, R. B. **Análise de Inadimplência Financeira na Área de Concessão de Crédito Rotativo usando RNA** – Goiânia: CPGEEC da UFG, 2001.
- [12] – GITMAN, L. J. **Princípios de Administração Financeira**. Editora Harbra. 7ª edição – São Paulo, 1997.
- [13] – GUIMARÃES, I. A. **Construção e Avaliação de uma Regra de Reconhecimento e Classificação de Clientes de uma Instituição Financeira com Base na Análise Multivariada**. Dissertação de Mestrado, UFPR, 2000.
- [14] – HARTMANN, F. R., **Redes Neurais, Conceitos Básicos e Análise**. Universidade do Vale do Rio dos Sinos, UNISINOS - RS, 2002.
- [15] – HAYKIN, Simon. **Redes Neurais: princípios e prática**. 2. ed. – Porto Alegre: Bookman, 2001. 900p.
- [16] – JOHNSON, R. A.; WICHERN, D. W. – **Applied Multivariate Statistical Analysis** – 4ª Ed. Prentice-Hall, inc., 1998.
- [17] – LATIF, S. A. **A Análise Fatorial auxiliando a resolução de um problema real de Pesquisa em Marketing**. Caderno de Pesquisa em Administração. São Paulo. USP, V.1, Nº 2, 2º sem./2000.

- [18] – LIMA, J. D. – **A Análise Econômico-Financeira de Empresas sob a ótica da estatística Multivariada**, Dissertação de Mestrado, UFPR – 2002.
- [19] – LIMA, J. D; MARQUES, J. M. **A Estatística Multivariada na Análise Econômico-Financeira de Empresas**. REVISTA FAE, Curitiba, v.5, n.3, p.59-59, set. /dez. 2002.
- [20] – MADALOZO, H. C. **Desenvolvimento de Ferramenta de Análise de Curvas Horizontais e Rodoviárias, para melhoramento de Projeto e Operações, Utilizando Redes Neurais**. Dissertação de Mestrado, UFPR, 2003.
- [21] – MALHOTRA, R; MALHOTRA D. K. **Evaluating Consumer Loans using Neural Networks – The International Journal of Management Science**, USA, 2003.
- [23] – SANTOS, J. O. dos. **Uma Contribuição ao Estudo de Fatores Sistemáticos Influenciadores da Inadimplência de Pessoas Físicas em Empréstimos Bancários** – Tese de Doutorado, São Paulo, EAESP-FGV, 2000.
- [24] – SCHRICKEL, W. K. **Análise de Crédito: Concessão e gerência de empréstimos**. 4ª. ed. São Paulo: Atlas, 1998.
- [25] – SERASA - <http://www.serasa.com.br/publicacao>, 2003.
- [26] – SICREDI - **Legislação e Formalística de Crédito** – Sistema Integrado de Crédito Cooperativo, SICREDI CENTRAL PARANÁ, 2003.
- [27] – SILVA, L. C. **Método robusto para a calibração de câmeras em estereofotogrametria**, Rio de Janeiro, UFRJ, 2003.
- [28] – SOUZA A. F., LUPORINI C. E. M. e SOUZA, M. S. de. **Caderno de Pesquisas em Administração**, São Paulo, v.1, nº 3, 1º trim. / 2000.

- [29] – STEINER, M. T. A.; CARNIERI, C.; KOPITTKKE, B.H; STEINER NETO, P. J. **Sistemas Especialistas Probabilísticos e Redes Neurais na Análise de Crédito Bancário**. Revista de Administração da USP (*RAUSP*), São Paulo. 34, n.3, p.56-67, 1999.
- [30] – STEINER, M. T. A.; ADAMOWICZ, E. C.; STEINER NETO, P. J. **Redes Neurais na Análise Econômica-Financeira de Empresas para Concessão de Crédito** – VI Congresso Brasileiro de Redes Neurais, São Paulo, 2003.
- [31] – WESTON, J. F. & BRIGHAM, E. F, **Managerial Finance**. New York: Holt, 1972
- [32] – WONG, B. K.; SELVI, Y. **Neural network applications in finance: A review and analysis of literature (1990-1996)**. Hong Kong Baptist University, Kowloon Tong, Hong Kong. Istanbul University, Turkey, 1998.
- [33] – YANG, B.; LI, L. X.; JI, H.; XU, J. **An early warning system for loan risk assessment using artificial neural networks**. Dong Hua University, Shanghai, China. Old Dominion University, Norfolk, USA. National Key Laboratory for Manufacturing Systems Engineering, Xian, China, 2001.

**ANEXOS**

## TABELAS COMPLEMENTARES

### INTRODUÇÃO

Durante a realização do presente trabalho, diversas manipulações foram realizadas. Várias tabelas intermediárias ou complementares foram geradas. Neste anexo serão apresentadas algumas tabelas que contribuíram para o andamento do processo, não colocadas nos capítulos originais com o objetivo de não saturar a apresentação. Desta forma, caso se deseje um detalhamento maior, cabe aqui verificar as tabelas apresentadas.

As variáveis informadas pelo(a) candidato(a) a cooperado são, de acordo com a cooperativa de crédito:

---

Conta:	Data Abert.:	Est. Civil:
Dependentes:	Profissão:	Residência:
Tempo:	(em anos)	
Informações Profissionais:	Tempo Serv:	(anos) Cargo:

---

### RELACAO DOS BENS DOS ASSOCIADOS – ANALITICO

---

Imóveis Rurais:	Área: (Há)	Vlr. Atual:
Aquis.:	Situacao:	Vlr. Atual:
Vlr. Rec:	CPF/CGC An.:	Endereço:
Bairro:	Cod. Munic.:	UF:

---

Cod. do Bem:	Descrição:	Descr. Compl:
MATRICULA:	Endereço:	Bairro:
Cod. Munic.:	UF:	CEP:

Area:	Aquis:	Reg. Imov:
Situacao:	Dt. Liber:	Vlr. Atual:
Vlr.Rec.:		

---

Imóveis Urbanos:	Cod. do Bem:	Descricao:
Endereço:	Bairro:	Cod. Munic.:
UF:	CEP:	Área:
Dt.Aquis:	Vlr. Atual:	Reg. Imov:

---

Maquinas e Equip. Agricolas:		Cod. do Bem:
Descrição:	Ano Fabric.:	Serie:
Chassi:	Modelo:	Percentual:
Vlr.Bem:	Vlr. Perc.:	

---

Veiculos:	Cod. do Bem:	Tipo:
Vlr. Atual:	Ano:	

---

Rebanhos e Semoventes:	Cod. do Bem:	Descrição:
Vlr. Atual:	Qtde:	Vlr. Total:

---

Produtividade:	Tipo - SOJA EM GRAOS	Area/Qtde:
Unid.:		

---

Seguros	Tipo:	Descrição:
Marca:	Vlr. Cobert.:	Vlr. Prêmio:

---

Despesas/Receitas:	Tipo:	Valor:
Rec./Disp.:	Periodicid:	

---

Resumo Financeiro do Associado:

---

Imóveis Rurais:

Imóveis Rurais - Benfeitorias:

Imóveis Urbanos:

Maquinas e Equipam. Agrícolas:

Veículos:

Rebanhos e Semoventes:

Outros Bens:

Participações em Empresas:

Capital Social (Conta Capital):

-----

Patrimônio Líquido:

-----

Receitas - Despesas Mensais :

Apl. Curto Prazo-Outros Bancos :

Apl. Curto Prazo-No Sistema :

(-) Obrigações de Curto Prazo :

(-) Obrig. Curto Prazo-Vencidas :

(-) Obrigações de Longo Prazo :

(-) Obrig. Longo Prazo-Vencidas :

(-) Carteira de Empréstimos :

(-) Adiantamento a Depositante :

(-) Limite de Cheque Especial :

-----

Situação Patrimonial :

-----

Saldo On-Line Conta Corrente:

Utilização de Cheque Especial:

-----

Associado:

CPF / CNPJ:

Risco:

SPC / SERASA:

Conta Corrente:

abertura (mês/ano):

Cheque Especial:

Limite:

média Mensal:

média Trimestral:

créditos no mês:

Cheques Devolvidos:

alínea (11)

alínea (12)

alínea (13)

Conta capital (Saldo):      Captação:      Saldo:

Aplicações a prazo:      Fundo investimento (Saldo):

existem títulos de cobrança para o associado?

O associado possui cartão de crédito?

Crédito;

Geral:      Nº Operações:      Valor:

Rural:      Nº Operações:      valor:

-----  
 Histórico dos últimos 12 meses:      Nº de parcelas pagas:

Parcelas pagas em atraso a partir de 15 dias

-----  
 Existem operações de crédito em inadimplência para o associado?

Existem operações de crédito a vencer para o associado em 30 dias?

-----

A maioria das variáveis citadas acima não foi aproveitada na realização deste trabalho. Os motivos variam desde a necessidade de se manter sigilo, o que é inclusive exigido por Lei, ao simples fato de que algumas variáveis não tem nenhum potencial como discriminantes, passando pela dificuldade em se atribuir valores numéricos a algumas delas e até mesmo não se contemplarem como variável comum ao conjunto e outras citadas no corpo do texto.

O Quadro abaixo refere-se ao perfil de um número de associados utilizados nas redes neurais desenvolvidas. Uma visão detalhada das variáveis pode ser observado no Quadro 3.4.

Coluna	Significado
01	Informações excludentes: SPC / SERASA;
02	Limite de crédito (cheque especial);
03	Média trimestral (conta corrente);
04	Conta capital (saldo);

Continua

Continuação

Coluna	Significado
05	Cheques devolvidos (alínea 11, 12, 13);
06	Cobrança (existência);
07	Histórico dos últimos 12 meses, referente a crédito (número de parcelas pagas);
08	Histórico dos últimos 12 meses, referente a crédito (parcelas pagas com atraso a partir de 15 dias);
09	Inadimplência (existem operações de crédito em inadimplência para o associado?).

Colunas								
01	02	03	04	05	06	07	08	09
0	300	91,91	1	492,60	0	0	0	0
0	0	2313,42	0	2137,06	0	10	1	0
0	1500	1880,00	0	1792,99	0	9	0	0
0	200	210,81	3	1283,64	1	6	2	0
0	400	228,35	0	1124,01	0	8	1	0
0	400	681,60	0	420,33	0	20	1	0
0	0	0,22	0	2010,93	0	10	1	0
0	400	60,07	0	1336,88	0	7	1	0
0	400	1672,60	0	740,56	0	7	1	0
0	2000	33,13	0	789,02	0	9	1	0
0	0	812,14	3	998,20	0	6	3	1
0	1500	21380,20	1	2003,33	0	8	0	0
0	500	521,08	0	199,21	0	6	1	0
0	100	996,21	2	838,67	1	10	4	0
0	1000	840,38	0	2063,31	0	7	1	0
0	800	413,50	1	522,06	0	2	0	0
0	0	878,64	0	3392,34	0	4	0	0
0	300	292,98	1	1738,68	0	4	0	0
0	300	849,11	4	791,58	1	9	6	0
0	200	19,98	0	1850,90	0	18	1	0

Continua

Continuação

0	800	158,71	0	2920,09	0	5	1	0
0	100	1251,04	6	592,99	1	10	6	0
0	200	231,22	0	1980,06	0	12	1	0
0	1000	11070,07	1	1092,78	0	12	1	0
0	500	704,18	0	2003,60	0	6	1	0
0	500	210,31	3	928,82	1	15	2	0
0	0	4858,70	1	2957,68	0	4	0	0
0	800	179,41	4	1018,53	0	8	3	1
0	200	672,02	0	1101,07	0	8	1	0
0	1000	3757,21	0	13110,79	0	1	1	0
0	600	139,32	2	860,10	0	9	3	1
0	700	749,03	1	1040,23	0	4	0	0
0	900	43,33	0	2954,23	0	10	1	0
0	100	568,10	3	1106,39	0	12	5	1
0	100	1102,60	1	2074,23	0	3	0	0
0	600	285,08	0	379,00	0	3	1	0
0	4000	607,24	0	1818,47	0	20	1	0
0	0	361,86	1	1980,00	0	6	0	0
0	500	802,07	0	2999,20	0	4	1	0
0	200	1021,15	4	1229,46	1	7	3	0
0	600	270,28	3	1169,66	0	10	4	1
0	1000	1456,42	0	1086,34	0	6	1	0
0	700	161,82	3	865,57	0	6	5	1
0	200	1321,31	3	976,23	1	12	5	1
0	900	103,46	2	781,82	0	12	6	1
0	400	22,92	1	1886,04	0	8	1	0
0	1000	22,91	0	1444,09	0	10	1	0
0	500	82,99	0	1813,31	0	8	1	0
0	500	786,60	0	5048,85	0	7	1	0
0	2000	8604,34	1	2208,55	0	3	0	0
0	800	290,23	1	2000,62	0	7	1	0
0	600	182,64	4	903,22	0	12	3	1
0	0	22,81	1	1488,85	0	8	0	0
0	500	502,91	0	1901,20	0	11	1	0
0	300	658,11	4	791,74	1	9	3	0
0	500	215,51	3	593,62	0	6	5	1
0	500	176,22	3	815,49	1	8	6	0
0	900	82,05	1	1885,83	0	6	1	0
0	900	2145,40	1	2656,61	0	12	0	0
0	400	737,52	1	1349,08	0	7	0	0
0	0	30,59	0	2568,39	0	18	1	0

Continua

Continuação

0	200	1606,72	3	635,21	1	6	5	1
0	100	807,09	0	1096,08	0	12	0	0
0	0	2200,13	0	2977,87	0	8	1	0
0	2000	276,50	0	1266,22	0	20	0	0
0	200	687,01	0	2817,27	0	13	1	0
0	300	37,79	1	1981,87	0	9	1	0
0	3000	21,10	1	2886,82	0	18	0	0
0	500	904,44	1	1000,09	0	6	0	0
0	200	224,44	1	410,00	0	5	1	0