



Universidade do Estado do Rio de Janeiro

Instituto Politécnico-IPRJ

Pós-Graduação em Modelagem Computacional

Previsão de inadimplência e redes neurais artificiais

Lourenço da Rocha Guimarães

Nova Friburgo, RJ – Brasil

Julho de 2006

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Catálogo-na-publicação

**Rede Sirius – Rede de Bibliotecas UERJ - Biblioteca CTC/E
(Instituto Politécnico)**



Universidade do Estado do Rio de Janeiro

Instituto Politécnico-IPRJ

Pós-Graduação em Modelagem Computacional

Previsão de inadimplência e redes neurais artificiais

Lourenço da Rocha Guimarães

Dissertação submetida ao Corpo Docente do Programa de Pós-Graduação em Modelagem Computacional do Instituto Politécnico da Universidade do Rio de Janeiro, como parte dos requisitos necessários para a obtenção do Grau de

Mestre em Modelagem Computacional

Linha de Pesquisa: Matemática Aplicada e Computação Científica

Dissertação aprovada por

Carlos Alberto Fialho Thompson Leite, Ph.D. – *orientador*
IPRJ / UERJ

Joaquim Teixeira de Assis, D.Sc.
IPRJ / UERJ

Washington Braga Filho, Ph.D.
PUC-RJ

Carlos Frederico Estrada Alves, D.Sc.
IPRJ/UERJ

Dedico este trabalho à minha esposa Ellen,
aos meus pais Laura e Jorge, e a toda a
minha família.

"Não espere por uma crise para descobrir o que é importante em sua vida."

Platão

AGRADECIMENTOS

Agradeço a Deus por me dar força e perseverança nos momentos mais difíceis.

À minha família pelo apoio, amor e compreensão pelas minhas ausências, durante a elaboração deste trabalho.

Ao meu orientador, Ph.D. Carlos A. Thompson , pela oportunidade e pelas orientações essenciais para a realização deste trabalho.

Aos amigos Antônio Merendáz (e a toda sua família), e Diego Roman, por todo o apoio.

Aos funcionários da Secretaria da PGMC-IPRJ/UERJ pela atenção e a sempre presente disposição de ajudar.

Aos funcionários da biblioteca do IPRJ, sempre dispostos a colaborar, especialmente pelo auxílio nas pesquisas bibliográficas, tão valiosas.

Aos colegas de trabalho, por todo o apoio dado.

A todos os companheiros que direta ou indiretamente contribuíram para a realização deste trabalho.

RESUMO

O objetivo deste trabalho é a previsão de inadimplência. Foram implementados dois modelos de previsão de inadimplência, de modo que o primeiro modelo fez uso de uma rede neural *feedforward* utilizando o algoritmo de retro propagação, e o segundo utilizou uma rede não supervisionada (rede Kohonen). As características relevantes de usuários de crédito foram apresentadas para as redes neurais, para o seu treinamento e teste. Os resultados obtidos demonstram que tanto as redes neurais supervisionadas quanto as redes neurais não supervisionadas mostraram-se instrumentos eficazes para o processo de previsão de inadimplência.

PALAVRAS-CHAVE: Redes Neurais Artificiais, Previsão de inadimplência, Kohonen

ABSTRACT

The objective of this work is the insolvency forecast. They were implemented two insolvency forecast models, so that the first model used a feed forward neural network utilizing the retro propagation algorithm, and the second utilized a non-supervised neural network (Kohonen networks). The prominent characteristics of credit users were presented for the neural networks, for their training and test. The results obtained showed that the supervised network as well the non-supervised neural network showed themselves efficient instruments for the insolvency forecast trial.

Key-Words: Neural Networks, Forecast of insolvency, Kohonen.

NOMENCLATURA

w	Peso sináptico
x	Valores apresentados aos neurônios da camada de entrada
b	Bias
a	Entrada para a função de transferência
y	Valores de saída de uma camada de uma rede neural supervisionada
d	Valores alvos (redes supervisionadas)
f	Função de transferência
$w_{ij}(t+1)$	Atualização dos pesos para uma iteração $(t+1)$, a partir de uma iteração
t	Iteração
\hat{W}	Neurônio vencedor
W	Neurônio apresentado à rede
L	Parcela de aprendizado que cada neurônio \hat{W} recebe em conjunto com sua vizinhança.
$\alpha(t)$	Taxa de aprendizagem
$\sigma(t)$	Desvio padrão
$V_m(t)$	Definição da vizinhança (rede não supervisionada)
D_j	Distância entre um determinado neurônio j e o neurônio vencedor
Ω	Número de épocas para treinamento
j	Neurônio
$e_j(n)$	Sinal de erro na saída do neurônio j
$d_j(n)$	Valor esperado de saída do neurônio j
$E(n)$	Valor instantâneo da energia total do erro

E_{med}	Energia média do erro quadrado
$v_j(n)$	Campo local induzido
m	Número total de entradas aplicadas ao neurônio j
$\varphi_j(v_j(n))$	Função de ativação
$\Delta w_{ji}(n)$	Correção aplicada ao peso sináptico $w_{ji}(n)$,
η	Parâmetro da taxa de aprendizagem do algoritmo da retropropagação
$\delta_i(n)$	Gradiente local
k	Neurônio de uma camada oculta
\bar{x}	Média aritmética
X	Variável de um conjunto de dados
σ_{XY}	Covariância entre X e Y
ρ	Coefficiente de correlação de Pearson
χ^2	Teste qui-quadrado
H_0	Hipótese nula
H_a	Hipótese alternativa

LISTA DE FIGURAS

		Página
Figura 2.1	Rede <i>feedforward</i> com três camadas	11
Figura 2.2	Função sigmóide	12
Figura 2.3	Arquitetura de uma rede SOM	15
Figura 2.4	Conceito de região de vizinhança em um arranjo retangular de neurônios	17
Figura 3.1	Gráfico comparativo de percentual de acertos para as quatro amostras	50
Figura 3.2	Gráfico comparativo de valores de χ^2 para as quatro amostras	50
Figura B.1	Teste do χ^2	69
Figura B.2	Tabela de valores do χ^2	68

LISTA DE TABELAS

	Página	
Tabela 3.1	Estado civil	23
Tabela 3.2	Tipo de emprego	23
Tabela 3.3	Faixa salarial	24
Tabela 3.4	Gênero	24
Tabela 3.5	Tipo de residência	24
Tabela 3.6	Tempo de residência	24
Tabela 3.7	Classificação numérica por estado civil	28
Tabela 3.8	Classificação numérica normalizada por estado civil	28
Tabela 3.9	Classificação numérica dos tipos de emprego	29
Tabela 3.10	Classificação numérica normalizada dos tipos de emprego	30
Tabela 3.11	Segunda classificação numérica por estado civil	30
Tabela 3.12	Segunda classificação numérica dos tipos de emprego	31
Tabela 3.13	Correlação entre os pares de atributos	32
Tabela 3.14	Exemplo de dados de uma amostra para treinamento	35
Tabela 3.15	Valores esperados de saída testados	36
Tabela 3.16	Valores de média e desvio padrão para os inadimplentes (amostra 1)	38
Tabela 3.17	Valores de média e desvio padrão para os adimplentes (amostra 1)	38
Tabela 3.18	Valores de média e desvio padrão para os inadimplentes (amostra 2)	38
Tabela 3.19	Valores de média e desvio padrão para os adimplentes (amostra 2)	39
Tabela 3.20	Valores de média e desvio padrão para os inadimplentes (amostra 3)	39
Tabela 3.21	Valores de média e desvio padrão para os adimplentes (amostra 3)	39

Tabela 3.22	Valores de média e desvio padrão para os inadimplentes (amostra 4)	40
Tabela 3.23	Valores de média e desvio padrão para os adimplentes (amostra 4)	40
Tabela 3.24	Valores médios de média e desvio padrão para os inadimplentes (das quatro amostras)	40
Tabela 3.25	Valores médios de média e desvio padrão para os adimplentes (das quatro amostras)	41
Tabela 3.26	Média de acertos e de χ^2 para a Amostra 1	42
Tabela 3.27	Média de acertos e de χ^2 para a Amostra 2	42
Tabela 3.28	Média de acertos e de χ^2 para a Amostra 3	42
Tabela 3.29	Média de acertos e de χ^2 para a Amostra 4	42
Tabela 3.30	Pesos obtidos após o treinamento com a amostra 1	46
Tabela 3.31	Pesos obtidos após o treinamento com a amostra 2	46
Tabela 3.32	Pesos obtidos após o treinamento com a amostra 3	46
Tabela 3.33	Pesos obtidos após o treinamento com a amostra 4	47
Tabela 3.34	Média de acertos e de χ^2 para a amostra 1 (treinamento não supervisionado)	47
Tabela 3.35	Média de acertos e de χ^2 para a amostra 2 (treinamento não supervisionado)	47
Tabela 3.36	Média de acertos e de χ^2 para a amostra 3 (treinamento não supervisionado)	48
Tabela 3.37	Média de acertos e de χ^2 para a amostra 4 (treinamento não supervisionado)	48
Tabela 3.38	Comparativo de percentual de acertos e valor de χ^2 para a amostra 1	49
Tabela 3.39	Comparativo de percentual de acertos e valor de χ^2 para a amostra 2	49
Tabela 3.40	Comparativo de percentual de acertos e valor de χ^2 para a amostra 3	49
Tabela 3.41	Comparativo de percentual de acertos e valor de χ^2 para a amostra 4	49

ÍNDICE

Capítulo:	Página:
Capítulo 1. Introdução	1
Capítulo 2. Fundamentação teórica	9
2.1 Tipos de arquiteturas das redes neurais	9
2.2 Tipo de aprendizagem	10
2.3 Rede neural supervisionada	11
2.4 O Algoritmo de retro propagação do erro	13
2.5 Redes SOM	13
2.6 Variáveis discriminantes/atributos	20
Capítulo 3. Experimentações numéricas e resultados	
3.1 Amostra	22
3.2 Variáveis discriminantes	23
3.3 Metodologia	25
3.4 Implementação dos modelos de Redes Neurais Artificiais	30
3.5 Resultados	46
Capítulo 4: Conclusões e trabalhos futuros	51
Referências bibliográficas	55
Apêndice	58

CAPÍTULO 1

INTRODUÇÃO:

O risco de crédito é a mais antiga forma de risco nos mercados financeiros. O risco de crédito é tão antigo quanto os empréstimos em si e a inadimplência é o principal fator que aflige qualquer instituição financeira. Afinal, a aprovação dos créditos e a definição das taxas a serem cobradas são decorrentes dos riscos que envolvem a carteira de operações (portfólios).

A estabilidade atingida pela economia brasileira levou bancos, e instituições financeiras em geral, a priorizar a concessão de crédito como forma de incrementar seus lucros. Após um período onde o mais importante era emprestar, chega-se a um período de maior maturidade onde as instituições percebem claramente que ceder crédito (emprestar) só é um bom negócio quando há a ADIMPLÊNCIA, ou seja, quando o beneficiário de crédito cumpre com a sua parte no acordo e efetiva o pagamento do que foi tomado como crédito, nas formas previstas quando da concessão deste. Fez-se então necessário que instituições financeiras, não só as brasileiras, mas em toda a economia mundial, desenvolvessem ou modificassem algum modelo de análise de crédito, de modo que fossem criados instrumentos capazes de auxiliar os executivos na tarefa de decidir sobre quando conceder ou não crédito, através da classificação do candidato ao crédito em “*provável adimplente*” e “*provável inadimplente*”. Cabe salientar que apesar de toda a evidência que vem sendo dada a este assunto atualmente, as pesquisas visando à criação de modelos de previsão de inadimplência já vêm acontecendo literalmente há décadas; Edward Altman apresentou em 1968 um trabalho propondo o uso do método de

discriminantes múltiplos (method of multiple discriminant analysis -MDA) para distinguir companhias como falidas e não-falidas (Altman, 1968).

Decisões ligadas a crédito remetem sempre à questão: haverá sucesso no recebimento do valor acordado quando da concessão do empréstimo ou não? Um modelo eficiente de análise de risco de crédito tem sempre que levar em conta estes fatores, maximizando a possibilidade do primeiro, e diminuindo o risco do segundo.

A concessão de crédito corporativo no Brasil é um importante instrumento para o desenvolvimento econômico, e é a principal atividade bancária. Segundo Silva (1988, p. 23):

“O crédito pode ainda fazer com que as empresas aumentem seus níveis de atividade, financiando a execução de projetos para os quais as empresas não disponham de recursos próprios suficientes”.

O fortalecimento e a solidez do sistema financeiro são essenciais para que haja recursos para o crédito, e em consequência, as empresas possam contar com esta alternativa para financiar suas atividades.

Nos últimos anos, em razão do crescimento do mercado de crédito no Brasil, o assunto passou a merecer grande atenção e preocupação, notadamente em razão da forte elevação índices de inadimplência.

A análise de crédito feita pelos bancos tem papel crucial para a sua sobrevivência, visto que, um crédito que seja concedido e que não retorne para o concedente, pode acarretar a perda do ganho obtido junto a várias outras operações.

Segundo Reed e Gill (1994), as operações de crédito inadimplidas causam forte impacto na liquidez bancária, gerando perdas, que, por sua vez, têm como

conseqüência a redução da capacidade de os bancos servirem seus clientes, bem como contribuir para o crescimento econômico da comunidade.

Segundo Caouette (2000), perdas com operações de crédito inadimplidas fizeram com que grandes instituições bancárias francesas e espanholas chegassem próximo à situação de falência, em anos recentes, e o valor envolvendo operações creditícias inadimplidas no sistema bancário japonês chega à casa dos US\$ 500 bilhões.

Conforme Altman e Haldeman (1995, p.11), “[...] a desconfiança sobre a consistência das pontuações de crédito subjetivas e um desejo por definições matemáticas para tais pontuações [...] têm gerado grande interesse em modelos objetivos e reprodutíveis”.

Segundo Back et al (1996), são duas as principais linhas de pesquisa nesta área: a primeira é a investigação empírica por indicadores que tornem possível uma classificação com menores taxas de erro. Esta linha é a que tem sido utilizada com maior freqüência. A segunda concentra-se na busca por melhores métodos estatísticos que possam ter como conseqüência o aperfeiçoamento da acurácia dos modelos de previsão.

Segundo Matarazzo (1998), é comum dois analistas de balanços chegarem a conclusões diferentes a respeito da concessão ou não de crédito para um determinado candidato a crédito. Logo, a existência de um processo decisório, sem se basear em avaliações e conceitos subjetivos do tomador de decisões, torna-se de fundamental importância. Cabe, então, o questionamento: as redes neurais artificiais, RNA, poderiam ser um instrumento útil neste tipo de processo decisório? No ponto em questão, abordado por Matarazzo, sem dúvida as RNA se apresentam como uma

opção real, visto que, após o devido treinamento, uma rede neural artificial é capaz de nos apresentar um resultado, para o estudo em questão, sem se basear nas avaliações e nos conceitos subjetivos de um analista de crédito.

As Redes Neurais Artificiais são sistemas computacionais de processamento paralelo tendo, segundo Braga (2000), seu funcionamento inspirado no cérebro humano. Segundo Haykin (2001, Prefácio), as RNA possuem “... a habilidade de aprender a partir de dados de entrada, com ou sem professor”.

Para Steiner et al. (1999), qualquer erro na decisão da concessão de crédito pode significar que em uma única operação haja perda do ganho obtido em dezenas de outras bem-sucedidas. Como a aprovação dos créditos e a definição das taxas de juros são decorrentes dos riscos que envolvem as carteiras de operações, e fica fácil entender a importância do assunto: A existência de crédito a uma taxa de juros aceitável proporciona o crescimento econômico de um país.

Caouette (2000, p. 24) coloca, sobre novas abordagens ao risco de crédito: “Fugindo da visão tradicional de que o julgamento de créditos é fundamentalmente uma “arte”, muitos bancos estão adotando novas abordagens. O Citibank, por exemplo, aplicou um sistema de inteligência artificial a um segmento de seu negócio de empréstimos comerciais”. Caouette (2000, p. 24) cita ainda, sobre o mesmo tema: “A década que se apresenta deve ser um período de fermentação, inovação e experimentação, à medida que novas abordagens sejam criadas, testadas e aplicadas”.

O professor do IBMEC, Duarte Jr afirma, em entrevista dada ao Jornal “O GLOBO” em 26/08/2005:

”Bancos continuam defasados em análise de crédito. A tarefa de melhorar a análise de crédito no Brasil é uma tarefa para os próximos 10 anos e resultará num melhor

apreçamento do crédito e em lucro maior para os bancos, já que as perdas serão reduzidas”.

Em novembro de 2005, foi manchete no caderno de economia do Jornal “O GLOBO”, o lucro recorde de R\$3,8 BI atingido até setembro daquele ano por um banco comercial. Sobre o assunto, o diretor-executivo de Controladoria do banco em questão, afirmou: “Tivemos um resultado sólido, fundamentado no crédito”. A reportagem também informava que a instituição comemorou o crescimento do crédito, mas, na outra ponta, começou a mostrar cautela com o risco de calote. O chamado índice de inadimplência (empréstimos vencidos há mais de 60 dias) pulou de 3,2% para 3,3% do total da carteira. Ainda de acordo com o balanço, as despesas com provisão para créditos de liquidação duvidosa cresceram 117,6%, chegando a R\$2,496 bilhões”. Na prática, isso diminuiu o lucro líquido da instituição no período.

O estudo de métodos de previsão de inadimplência se coloca como assunto de importância elevada, devido à relevância econômica e social. O uso de redes neurais artificiais para este fim se apresenta como uma possibilidade a ser estudada, devido a expectativa existente no que diz respeito à criação, teste e implementação de novos e mais objetivos métodos de previsão de inadimplência.

JUSTIFICATIVA E IMPORTÂNCIA DO TRABALHO

O crédito a pessoas físicas, nas diversas modalidades financeiras de cartão de crédito, crédito direto ao consumidor e cheque especial, tem sido tratado empiricamente ou por sistemas de *credit score* baseados em análise discriminante ou de regressão logística. No entanto, nos últimos anos, em razão do crescimento do mercado de crédito ao varejo no Brasil, o assunto passou a merecer maior atenção e preocupação, notadamente em razão da forte elevação das taxas de inadimplência, e

também, do interesse de instituições financeiras em desenvolver modelos de previsão de inadimplência que não recusem inadequadamente a concessão de crédito a clientes que irão honrar as operações de crédito compromissadas. A avaliação equivocada de um bom candidato a crédito tem como consequência a geração de ganhos não realizados ao se negar um crédito a um bom pagador, enquanto que, a concessão de crédito a um cliente erroneamente classificado como bom candidato tem como consequência provável o prejuízo pela perda do valor envolvido na operação possivelmente inadimplida. Um instrumento capaz de fazer essa classificação dos candidatos a crédito tem relevante importância para o sucesso da política de crédito de uma instituição financeira.

A decisão pela concessão de crédito aos clientes e por mudanças no perfil comportamental dos tomadores de crédito, decorrente de mudanças nas variáveis econômicas: alguém que não seria inadimplente sob condições normais, poderá vir a sê-lo sob redução da liquidez da economia, por exemplo. Os prejuízos causados às instituições financeiras que concedem o crédito inadimplido são enormes, e qualquer redução do risco poderá implicar em grandes benefícios.

Na maioria dos ambientes de avaliação de crédito, os gerentes trabalham com uma enorme variedade de informações vindas das mais diversas fontes. Muitas destas informações podem ser incompletas, ambíguas, parcialmente incorretas ou de relevância dúbia. Os gerentes analisam estas informações de maneira muito subjetiva e muitas vezes não conseguem explicar os processos de tomada de decisões. As Redes Neurais Artificiais oferecem uma eficiente alternativa para o tratamento de informações dessa natureza, uma vez que podem ser modificadas mais facilmente que sistemas convencionais, são de utilização rápida, e são mais simples que os

métodos convencionais na modelagem não linear, o que lhes dá uma grande vantagem sobre vários métodos estatísticos utilizados em economia e finanças, que são lineares (Trippi, 1993).

A área financeira vem demonstrando grande interesse em aplicações da Inteligência Computacional para o tratamento destes problemas (Apté, 1982). Vários projetos de inteligência artificial para utilização em sistemas de cartão de crédito vêm sendo implementados (Carter, 1987). Um sistema inteligente para auxiliar a decisão no tocante à autorização de compra com cartão de crédito utilizado pela *American Express* teve como consequência uma economia de 20% no tempo médio gasto em cada transação, e a qualidade das decisões tomadas por este sistema foram em média 50% melhores que decisões tomadas funcionários do setor de autorização de crédito que não utilizaram este instrumento (Treleaven, 1992).

OBJETIVO

O objetivo deste trabalho é a previsão de inadimplência. Foram implementados dois modelos de previsão de inadimplência, de modo que o primeiro modelo fez uso de uma rede neural supervisionada utilizando o algoritmo de retropropagação do erro, e o segundo utilizou uma rede não supervisionada (rede Kohonen).

ESTRUTURA DO TRABALHO

O Capítulo 1 apresenta a origem, importância, objetivos e estrutura do trabalho.

No Capítulo 2 são apresentadas informações sobre as redes neurais artificiais implementadas; conceitos e fundamentação teórica são abordados. Também são apresentadas informações sobre conceitos básicos de análise de crédito.

No Capítulo 3, são informados os experimentos numéricos e tratamentos dados às informações utilizadas para a construção e treinamento das redes neurais, assim com os resultados do trabalho. Constam neste capítulo informações sobre a metodologia, ou seja, são apresentados os dados referentes à construção dos modelos de previsão de adimplência baseados no uso das redes neurais, explanando a forma como as amostras para treinamento e testes foram tratadas/inseridas nos modelos acima referidos. Os resultados e suas análises também são apresentados no capítulo 3.

No Capítulo 4 são apresentadas as conclusões e as sugestões de futuros estudos, visando à continuação da pesquisa apresentada nesta tese.

Após o capítulo 4, são apresentadas as referências bibliográficas utilizadas no desenvolvimento do presente trabalho.

No apêndice, apresentam-se o algoritmo de retropropagação do erro utilizado nas redes neurais supervisionadas, e as informações sobre as técnicas e os fundamentos de estatística utilizadas no desenvolvimento do presente trabalho.

CAPÍTULO 2

FUNDAMENTAÇÃO TEÓRICA

2.1 TIPOS DE ARQUITETURAS DAS REDES NEURAIIS

A arquitetura de uma rede neural artificial a ser implementada está diretamente ligada com o tipo de problema a ser tratado. Segundo Braga (2000) a arquitetura de uma rede neural artificial restringe o tipo de problema a ser tratado pela rede em questão.

A arquitetura de uma RNA deve ser estabelecida de acordo com os seguintes parâmetros:

- Número de camadas da RNA;
- Número de neurônios em cada camada da rede;
- Conexão entre os neurônios.

No tocante ao número de camadas, as redes podem ser:

- 1) De camada única, onde existe, somente, uma unidade de processamento entre as entradas e as saídas da rede.
- 2) De múltiplas camadas: há várias camadas de neurônio entre as entradas e saídas da rede.

Algumas definições são apresentadas a seguir:

- **Camada de entrada:** unidade onde os padrões são apresentados à rede e passados para a(s) camada(s) subsequente(s);
- **Camadas ocultas ou intermediárias:** unidades onde os processamentos ocorrem, através das funções de ativação e atualização dos pesos sinápticos.

- **Camada de saída:** unidade onde o resultado final da rede é apresentado e comparado ao resultado desejado.

2.2 TIPOS DE APRENDIZAGEM

Segundo Haykin (2001, p. 75) “Há atividades demais associadas à noção de “aprendizagem” para justificar a sua definição de forma precisa”. Ele afirma, ainda, na mesma página: “Aprendizagem é o processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pela qual a modificação dos parâmetros ocorre”.

Os diversos modelos de redes neurais utilizam diferentes regras de treinamento, onde os pesos de suas conexões são ajustados de acordo com os padrões apresentados. “Um conjunto de regras pré-estabelecidas e bem definidas para a solução de um problema de aprendizagem é denominado algoritmo de aprendizagem” (Haykin, 2001, p.76).

Dentre os diversos tipos de algoritmo de aprendizagem existentes, os mais utilizados são:

- **Aprendizado supervisionado:** a resposta desejada é indicada à rede por um supervisor externo; de acordo com Braga (2000, p.16), “O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos”.

- **Aprendizado não-supervisionado:** “Não há exemplos rotulados da função a ser aprendida pela rede” (Haykin, 2001, p. 89). Não é fornecida à rede qualquer informação sobre uma resposta desejada para cada neurônio de saída, não havendo

portanto um professor ou supervisor externo para acompanhar o processo de aprendizagem (Braga, 2000).

2.3 REDE NEURAL SUPERVISIONADA

A rede neural utilizada é uma rede neural *feedforward*, multicamada, treinada pelo algoritmo *backpropagation*, utilizando como função de transferência a função sigmóide.

As RNA de múltiplas camadas têm sido aplicadas com sucesso para resolver diversos problemas difíceis, através de seu treinamento supervisionado, com o algoritmo de *retropropagação do erro* (Haykin, 2001) e a função sigmóide pode aproximar qualquer função contínua com derivada contínua sobre todo intervalo real (a função é diferenciável em qualquer ponto). A função sigmóide é a função de ativação mais comum utilizada na construção de redes neurais (Haykin, 2001).

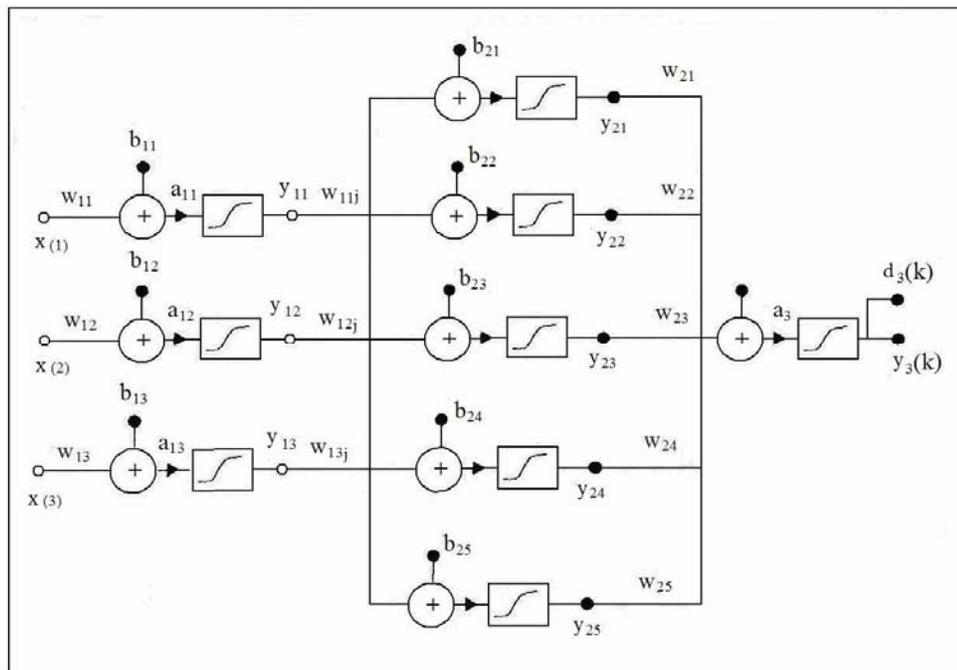
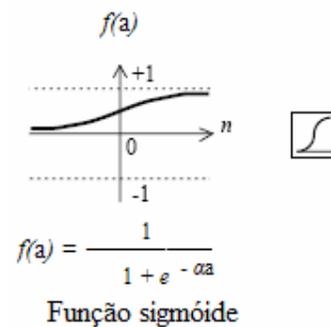


Figura 2.1 Rede *feedforward* com três camadas



**Figura 2.2: Função sigmóide (adaptada de
Neural Network Toolbox User's Guide)**

O algoritmo de retropropagação do erro utiliza pares (entrada, valor esperado de saída), e, através da correção dos erros, promove o ajuste dos pesos da rede (Braga, 2000). O objetivo do uso deste algoritmo é minimizar o erro entre o valor de saída da rede e o resultado desejado. Este algoritmo de aprendizado supervisionado visa modificar uma rede neural *feedforward*, pela minimização da função de erro, obtida pela comparação entre a saída da rede ($y(w)$) e o valor desejado de saída (d). Os pesos são modificados a cada época visando diminuir o valor da função de erro nas épocas subseqüentes de treinamento, conforme os passos abaixo:

Passo 1 (*feedforward*): os dados dos elementos da amostra de treinamento (entradas) são apresentados, e seguem o fluxo de ativações até que eles alcançam à camada de saída.

Passo 2 (*backward*): os valores de erro são calculados para cada unidade da camada de saída, por comparar ao valor real apresentado (y , calculado na etapa 1) com o valor desejado de saída (d). Então, estes erros fluem da camada de saída às camadas

escondidas e à camada de entrada. Os valores de erro são usados para ajustar os pesos na camada de saída. Este treinamento, em geral, requer muitas épocas.

2.4 O ALGORITMO DE RETROPROPAGAÇÃO DO ERRO (BACKPROPAGATION ALGORITHM):

O resumo do algoritmo *backpropation* se encontra em Haykin (2001- pág. 188 à 193), e é apresentado no Apêndice 1, com pequenas adaptações da parte textual.

2.5 – REDES SOM (Self-Organizing Map)

O estudo da neurofisiologia permitiu que surgissem proposições de arquiteturas que simulassem o funcionamento do cérebro na armazenagem de conhecimento e no desempenho cognitivo. Estas redes são conhecidas como mapas auto-organizáveis de Kohonen ou redes SOM (Self-Organizing Map), as quais são descritas a seguir.

Segundo Haykin (2001), o mapa auto-organizável de Kohonen “...captura as características essenciais dos mapas computacionais do cérebro e ainda se mantém tratável do ponto de vista computacional”.

As redes SOM pertencem a um grupo de redes neurais artificiais chamado *redes competitivas* (Fausett, 1994). Estas redes combinam competição com uma forma de aprendizagem para efetuar os ajustes de seus pesos, fazendo uso de um treinamento não supervisionado. A rede busca encontrar similaridades baseando-se apenas nos padrões de entrada, tendo como principal objetivo o agrupamento dos

dados de entrada que são semelhantes entre si formando classes ou agrupamentos denominados *clusters*.

Padrões contidos em um conjunto de dados n-dimensional podem possuir uma estrutura bastante complexa. Segundo Haykin (2001), as redes SOM transformam um sinal incidente de dimensão arbitrária em um mapa discreto uni ou bi-dimensional, realizando esta transformação adaptativamente de forma topologicamente organizada. Os modelos SOM apresentam-se como uma forma de detectar similaridades relevantes em grandes bases de dados, fazendo uso da construção de um mapa facilmente caracterizável a partir de um espaço multidimensional, tomando assim os *clusters* mais facilmente identificáveis.

Como acontece com os modelos multicamadas genéricos de redes neurais, as redes SOM utilizam da adaptação dos pesos durante o processo de treinamento, mas este aprendizado não é ocorre de forma supervisionado, pois não é apresentado “a rede um valor de saída esperado.

Em uma rede SOM tem-se uma unidade de entrada para cada componente do vetor de entrada. Cada unidade (neurônio) de saída representa um *cluster*, limitando assim a quantidade de *clusters* ao número de saídas. Durante o treinamento da rede, o neurônio de saída que melhor responde ao vetor de entrada (isto é, apresentar a menor distancia euclidiana em relação ao vetor de entrada) será considerado o neurônio vencedor; os vetores de pesos para a unidade vencedora e para os neurônios vizinhos são ajustados de acordo com o algoritmo de treinamento a ser descrito na próxima seção.

A Figura 2.1 mostra uma arquitetura de uma rede SOM bidimensional com nove neurônios.

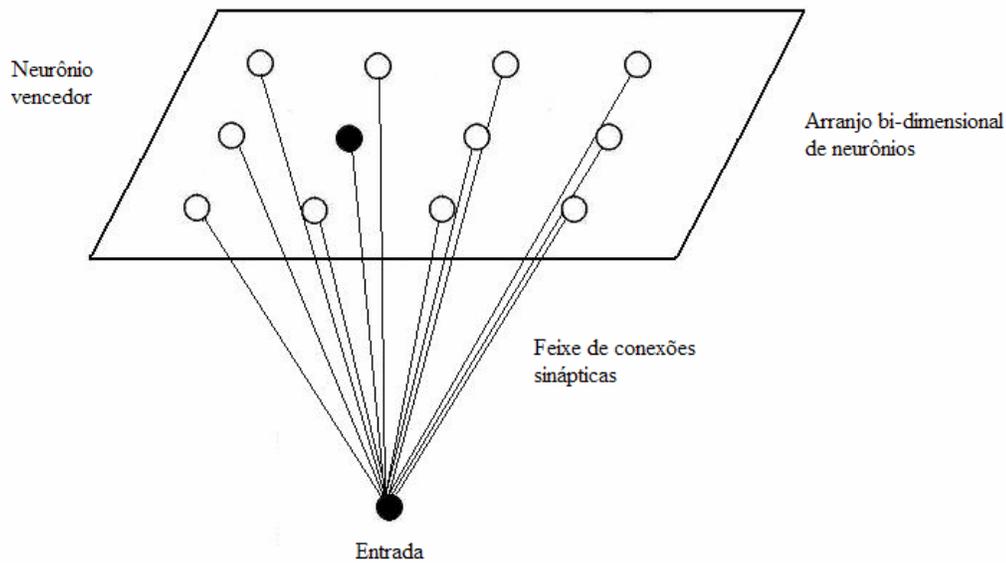


Figura 2.3: Arquitetura de uma rede SOM. Adaptado de Haykin (2001).

O algoritmo para a formação do mapa auto-organizável tem começo com a inicialização dos pesos sinápticos da grade (Haykin, 2001). Atribuem-se valores aleatórios aos pesos iniciais, de modo que não haja qualquer organização prévia imposta ao mapa de características.

Segundo Haykin (2001), os três processos a seguir são essenciais na formação do mapa auto-organizável:

- 1- Competição: A cada apresentação de um padrão de entrada, os neurônios competem entre si para que seja então determinado o neurônio vencedor, isto é, o neurônio que apresenta a maior similaridade com o padrão de entrada apresentado.
- 2- Cooperação: A localização espacial de uma região topológica é determinada, tendo como referência o neurônio vencedor.

- 3- Adaptação sináptica: Os pesos do neurônio vencedor e dos neurônios vizinhos sofrem alterações, de modo que suas respostas a exposições subseqüentes de um padrão de entrada similar sejam melhoradas.

Conceito de vizinhança:

Uma característica fundamental das redes SOM é que não somente o peso do neurônio vencedor é corrigido, mas também sofrem ajustes os pesos dos neurônios vizinhos, em um comportamento que visa fazer com que a rede simule uma atividade cerebral. Conforme Tafner (1996, pág. 124), “Com relação a este tema, vale lembrar que alguns cientistas, rastreando a atividade cerebral, notaram que, apesar de todas as células nervosas estarem ligadas e coligadas entre si, existem centros de atividade mais intensificados conforme a atividade exercida no momento”.

As redes SOM, de modo a simular o comportamento cognitivo do cérebro humano, efetuam os ajustes dos pesos em um processo onde o neurônio vencedor captura a maior parte do conhecimento adquirido acerca de determinados padrões apresentados, mas os pesos dos neurônios vizinhos ao neurônio vencedor também aprendem com a entrada apresentada em menor escala.

Em uma rede que possua apenas uma linha de neurônios, a região de vizinhança (onde se localizam os neurônios vizinhos ao vencedor) é composta pelos neurônios à direita e à esquerda do neurônio vencedor. Porém, quando lidamos com estruturas mais complexas, como por exemplo, em uma rede com estrutura bi-

dimensional de neurônios, outras possibilidades de região de vizinhança se apresentam, como pode se perceber através da análise da Figura 2.2.

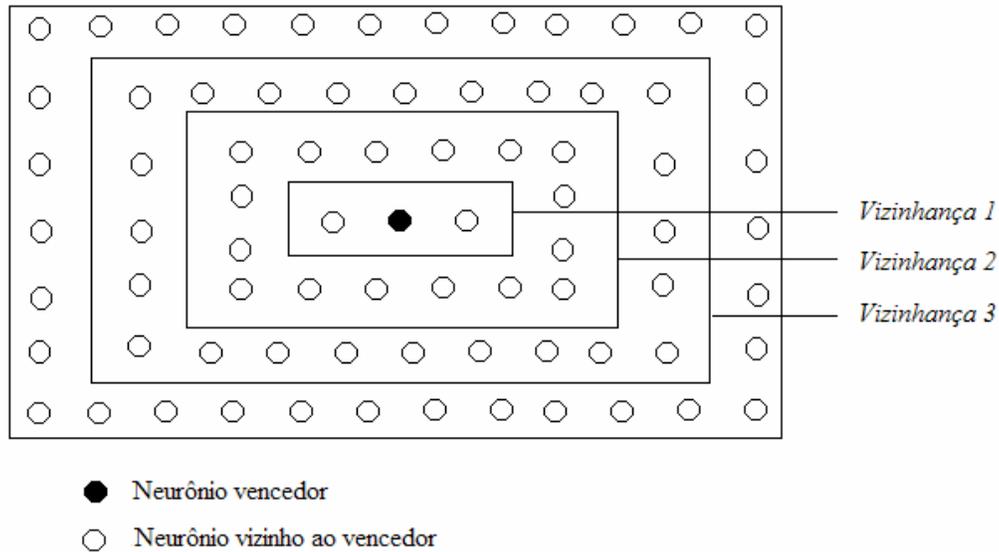


Figura 2.4: Conceito de região de vizinhança em um arranjo retangular de neurônios. Adaptado de Fausett, 1994.

Conforme Tafner (1996, pág. 127), “A melhor auto-organização, conforme observações do próprio Kohonen, ocorre quando o conjunto de vizinhos começa extenso e diminui com o tempo, podendo, então, a vizinhança ser definida como tempo variável”.

Inicialmente, o conjunto de vizinhos ao neurônio vencedor (região de vizinhança) pode ser extenso, de modo que haja uma especialização regional dos neurônios. Com o decorrer do processo de aprendizagem, a região de vizinhança deve diminuir gradativamente, tornando a resposta da rede mais localizada após cada iteração.

De maneira similar ao que ocorre com outros modelos de RNA, o algoritmo de aprendizado de uma rede SOM se baseia na apresentação de padrões de entrada, cálculo dos valores apresentados como saída dos neurônios, e no ajuste dos pesos, mas apresenta diferenças, que são:

- O cálculo dos valores de saída é feito com base na similaridade entre os valores dos pesos e dos atributos do padrão de entrada.
- O conceito de vizinhança, que determina que os pesos de um neurônio (o vencedor) sofram uma correção em maior escala do que os pesos dos demais neurônios, mas, ainda assim, permite que os pesos dos neurônios da região de vizinhança também sejam corrigidos, mesmo que em menor grau.

Operação e algoritmo de uma rede SOM

Conforme Tafner (1996), para que seja calculado o novo valor de um peso de um neurônio, devemos seguir a equação (2.14):

$$w_i(t+1) = w_i(t) + L(t)[e(t) - w_i(t)] \quad (2.14)$$

Uma vez definida a região de vizinhança em torno do neurônio vencedor W em uma iteração t , a parcela de aprendizado a ser recebida por cada neurônio \hat{W} da região de vizinhança está ligada à sua distância física (em número de neurônios) em relação ao neurônio vencedor W , e é determinada pela equação (2.15)

$$L = \alpha(t) \exp\left(\frac{-\|W - \hat{W}\|}{\sigma^2(t)}\right) \quad (2.15)$$

$W - \hat{W}$ é a distância física (expressa em quantidade de neurônios) entre o neurônio vencedor \hat{W} e o neurônio W pertencente à região de vizinhança.

A correção do valor do peso de cada neurônio também está ligada às funções $\alpha(t)$ e $\sigma^2(t)$, e, segundo Haykin (2001), estas funções devem diminuir com o tempo, isto é, a cada nova iteração termos novos valores para $\alpha(t)$ e $\sigma^2(t)$. Neurônios que não pertençam à região de vizinhança não terão seus pesos alterados na iteração em questão.

O algoritmo da rede SOM, segundo Tafner (1996), apresenta-se a seguir:

01. Especificar o número de iterações
02. Inicializar os pesos de todos os neurônios da rede com valores aleatórios.
03. Determinar o raio da vizinhança V_m de cada neurônio.
04. Apresentar uma entrada à rede.
05. Calcular a distância euclidiana entre a entrada e os pesos para cada neurônio da rede:

$$D_j = \|x - w_j\| = \sqrt{\sum_{i=1}^n (x_i - w_{ij})^2} \quad (2.16)$$

06. Selecionar o neurônio vencedor, isto é, o neurônio que apresentar a menor distância euclidiana.
07. Atualizar os pesos do neurônio vencedor e dos neurônios pertencentes à região de vizinhança definida no passo 3.

$$w_{ij}(t+1) = w_{ij}(t) + L[x_i - w_{ij}(t)], \quad (2.17)$$

08. Reduzir os valores de $\alpha(t)$ e $\sigma^2(t)$;
09. Continue a partir do passo 4 por Ω iterações.

A escolha de um determinado algoritmo de aprendizagem está intimamente relacionada com a tarefa que a rede deverá realizar. As redes SOM têm sido predominantemente utilizadas para agrupamento de padrões, capturando as

características importantes contidas em espaço de entrada multidimensional e desta forma, representá-lo num espaço bidimensional (Haykin, 2001).

2.6 VARIÁVEIS DISCRIMINANTES/ATRIBUTOS

Uma das decisões mais importantes na construção de uma rede neural é a escolha das variáveis nas quais a rede irá se basear para poder aprender. Haykin (2001, p 433) cita: “Um problema comum em reconhecimento estatístico de padrões é a seleção de características ou extração de características”. Quando as redes neurais são utilizadas para realizar previsão financeira, dados sobre diversos indicadores econômicos são utilizados como valores de entrada da rede para que na saída obtenha-se alguma informação, como neste caso, o indicativo se um candidato a crédito é um provável adimplente ou provável inadimplente.

Segundo Haykin (2001), “A seleção de características se refere a um processo no qual um espaço de dados é transformado em um espaço de características...”. Haykin (2001) cita ainda “Entretanto, a transformação é projetada de tal forma que o conjunto de dados pode ser representado por um número reduzido de características “efetivas” e ainda reter a maioria do conteúdo de informações intrínseco dos dados...”.

No tocante à análise de crédito, uma boa técnica discriminante deve ser capaz de permitir a construção de um modelo preditor de inadimplência com alto poder discriminador, permitindo a classificação de novos aspirantes a crédito. Caouette (2000) cita que cada característica do candidato a crédito deve ser analisada de forma a levar a construção de um sistema de escalas capaz de separar os bons candidatos a crédito dos maus candidatos a crédito, mas que também torne o processo de análise da concessão do crédito menos complicado. Segundo Securato (2002), deve ser

construído um sistema de pontuação que atribua pesos a cada parâmetro de acordo com a sua relevância.

O pré-processamento dos atributos dos elementos da amostra de treinamento tem ainda outra etapa que é a verificação da correlação e da covariância destas variáveis.

Segundo Gujarati (1995), caso o módulo do coeficiente de correlação entre dois atributos seja maior que 0,8, um destes atributos pode ser descartado.

Outro passo necessário no pré-processamento dos dados a serem apresentados à RNA é a normalização de seus valores. Segundo Haykin (2001, p. 2085), “Cada variável de entrada deve ser pré-processada, de modo que seu valor médio, calculado sobre todo o conjunto de treinamento, ou seja próximo de zero, ou seja pequeno comparado com o desvio padrão (LeCun, 1993)”. Para atingir a esse objetivo, as variáveis de entrada devem ser processadas de modo a ter seus valores dentro do intervalo [0,1].

No próximo capítulo serão abordadas as etapas de seleção das amostras para treinamento e teste das redes neurais, assim como todos os passos para a implementação dos modelos de RNA utilizados neste estudo. Serão relatados, também no próximo capítulo, os resultados obtidos.

CAPÍTULO 3

EXPERIMENTAÇÕES NUMÉRICAS E RESULTADOS

3.1 AMOSTRA

Para a condução deste trabalho, que consiste na classificação preventiva de pessoas físicas candidatas a crédito como “provável adimplente” ou “provável inadimplente”, foram empregadas como base quatro amostras de cento e trinta clientes de uma carteira de clientes de uma instituição financeira que opera com diversas linhas de crédito para pessoas físicas. Cada uma das quatro amostras utilizadas para o treinamento da rede neural consiste em um total de 100 tomadores de crédito (50 % com histórico de adimplência e 50 % com histórico de inadimplência) enquanto outros 30 tomadores de crédito (15 adimplentes e 15 inadimplentes) foram usados para a etapa de testes. A carteira de clientes de onde foram capturadas as informações das amostras de treinamento e de teste em questão é formada por um total de 4350 clientes com operações de crédito já contratadas, possibilitando, então, o conhecimento de um histórico real de adimplência ou de inadimplência para cada cliente da carteira. A seleção dos clientes que fizeram parte das amostras de treinamento e de teste foi feita de forma aleatória, e, depois da seleção, verificou-se que a amostra possuía clientes representantes de todas as categorias, isto é, havia clientes com os diversos estados civis, diversas faixas de renda, de diversos tipos de emprego, etc.

Com os princípios definidos para a coleta de dados, partiu-se então para a seleção das variáveis discriminantes.

3.2 VARIÁVEIS DISCRIMINANTES

O uso dos indicadores empregados como variáveis foi fruto de uma coletânea da literatura, seguida de uma verificação do poder discriminante de cada uma delas, exposta a seguir.

Após uma avaliação da relação adimplentes/inadimplentes de diversas características observáveis nos clientes que já possuíam histórico de operações de crédito, foram escolhidos os atributos que apresentaram diferença significativa entre as razões adimplentes/inadimplentes (possuíam maior poder discriminante na amostra); foram eles: renda, estado civil e tipo de emprego. As tabelas abaixo nos dão uma visualização da razão entre adimplentes e inadimplentes, no tocante aos três atributos mencionados acima.

Tabela 3.1: estado civil

Estado civil	Adimplentes	Inadimplentes	Razão
Separados	28	44	0,636
Solteiros	64	84	0,762
Casados	76	64	1,188
Viúvos	32	8	4,000

Tabela 3.2: tipo de emprego

Tipo de emprego	Adimplentes	Inadimplentes	Razão
Privado	8	108	0,074
S/ vínculo de emprego	16	32	0,500
Público	176	60	2,933

Tabela 3.3: faixa salarial

Faixa salarial	Adimplentes	Inadimplentes	Razão
<260,00	16	56	0,286
261,00 a 1000,00	24	60	0,400
1001,00 a 2500,00	128	68	1,882
2501,00 a 5000,00	32	16	2,000

Outras características de pessoas candidatas a crédito também foram analisadas no intuito de verificar seu poder discriminante. Assim sendo, os atributos gênero, tipo de residência e tempo de residência dos tomadores de crédito presentes na amostra em questão tiveram seu poder discriminante analisado conforme as tabelas abaixo:

Tabela 3.4: gênero

Gênero	Adimplentes	Inadimplentes	Razão
Masculino	92	88	1,045
Feminino	108	112	0,964

Tabela 3.5: tipo de residência

Tipo de resid.	Adimplentes	Inadimplentes	Razão
Própria	44	40	1,100
Alugada	92	104	0,885
Familiar	44	40	1,100
Cedida	20	16	1,25

Tabela 3.6: tempo de residência

Tempo de resid	Adimplentes	Inadimplentes	Razão
< um ano	40	32	1,250
Entre 1 e 5 anos	76	88	0,864
Entre 5 e 10 anos	52	56	0,929
> 10 anos	32	24	1,333

É possível perceber que, no tocante a esses três últimos atributos (gênero, tempo de residência e tipo de residência), não há forte predomínio de histórico de inadimplência ou adimplência entre as diversas classificações existentes entre cada uma dessas variáveis discriminantes. Ou seja, as características gênero, tipo de residência e tempo de residência **não** apresentam grande poder discriminante.

A verificação do poder discriminante de cada uma das características dos tomadores de crédito, através do cálculo da razão adimplentes/inadimplentes, permitiu que alguns conceitos popularmente aceitos como verdadeiros fossem repensados, sobretudo para a amostra em questão. As idéias de que “cliente com baixa renda costumam honrar seus compromissos ligados a crédito de forma mais assídua do que clientes de maior renda” e de que “mulheres pagam suas dívidas com maior pontualidade do que homens” não expressam a verdade, para a amostra utilizada para este estudo.

Chega-se a uma etapa em que se encontra a presença de variáveis categóricas (não métricas), que, portanto, exigem sua transformação em dados numéricos. São exemplos deste tipo de característica: gênero, tipo emprego, estado civil. Utilizamos a técnica *dummy variables* (variáveis fictícias) para acomodar as variáveis não métricas. Essa transformação dos dados não-métricos em dados métricos será abordada na seção 3.3.2 Pré-processamento dos dados.

3.3 Metodologia

A metodologia utilizada nas diferentes fases deste projeto incluiu várias estratégias para tratamento das informações das bases de dados, seleção, treinamento e testes das redes, como serão discutidas nas próximas sub-seções.

3.3.1 Coleta de Dados

Nesta primeira fase foram reunidos todos os dados pertinentes e potencialmente úteis à tarefa. Estes estavam contidos nas bases de dados fornecidas por uma instituição financeira, instituição esta que terá seu nome omitido deste trabalho, a pedido da mesma, com o intuito de preservação das informações ligadas à sua carteira de crédito, bem como das características de sua política creditícia. Os dados coletados incluíam informações pessoais dos clientes, informações sobre a aplicação de crédito solicitada e um histórico sobre a adimplência destes clientes, permitindo a classificação destes em duas classes: adimplentes e inadimplentes.

3.3.2 Pré-processamento dos dados

A presença de variáveis discriminantes categóricas (não métricas) exige transformação. Gênero, tipo de emprego, estado civil, área de atuação são exemplos desse tipo de informação a ser transformada. Foi utilizada a técnica *dummy variables* para a transformação das variáveis não métricas, de modo que dados simbólicos como estado civil e tipo de emprego foram transformados em classes numéricas, mais adequadas para utilização neste tipo de rede. Foram, então, atribuídos valores que para cada uma das características do candidato a crédito, valores estes proporcionais ao quanto esta variável discriminante estava associada a clientes com histórico de adimplência ou inadimplência. As tabelas 3.1 e 3.2 auxiliaram na visualização do poder discriminante de cada uma das características dos tomadores de crédito, e na posterior transformação destes dados não-métricos.

Seguindo o informado por Caouette (2000), construiu-se de um sistema de escalas, com o intuito de separar os bons candidatos a crédito dos maus candidatos a crédito. Seguindo o informado por Securato (2002), o sistema construído um sistema de pontuação que atribuiu pesos a cada parâmetro de acordo com a sua relevância.

No tocante ao estado civil, após a análise da tabela 3.1, fica claro que, para esta amostra, a razão adimplentes/ inadimplentes tem menor valor para o estado civil *separado*, aumentando então quando avaliamos os clientes solteiros. Os casados possuem a segunda melhor razão adimplentes/inadimplentes, enquanto o estado civil *viúvo* apresenta a melhor pontuação neste quesito. Cada estado civil foi representado por um número inteiro, entre o intervalo [1,4], conforme o seu poder discriminante, isto é, quanto mais ligado a histórico de inadimplência, menor seria a sua representação numérica; por outro lado, caso o estado civil em questão estivesse mais

ligado a um histórico de adimplência, maior seria o valor de sua representação numérica, conforme apresentado na tabela 3.7:

Tabela 3.7: Classificação numérica por estado civil

Estado civil	Classificação numérica
Separados	1
Solteiros	2
Casados	3
Viúvos	4

Como as os atributos a serem informados como entrada para as redes neurais precisavam estar dentro do intervalo $[0,1]$, providenciou-se a normalização da representação numérica dos diversos estados civis, dividindo cada representação em questão por 4, chegando-se aos resultados apresentados na tabela 3.8:

Tabela 3.8: Classificação numérica normalizada por estado civil

Estado civil	Classificação numérica normalizada
Separados	0.25
Solteiros	0.50
Casados	0.75
Viúvos	1.00

De forma similar, construiu-se a representação numérica de outro atributo não métrico, o tipo de emprego. O tipo de emprego, que originalmente não era um atributo numérico, também passa por uma transformação análoga à anterior, para que possamos então apresentar à rede este atributo dos candidatos a crédito.

Quando tratamos o atributo *tipo de emprego*, após a análise da tabela 3.2, chegamos à conclusão que, para esta amostra, a razão adimplentes/ inadimplentes tem menor valor para o emprego privado, aumentando então quando avaliamos os clientes sem vínculo de emprego. Os clientes que possuem emprego público ou são aposentados apresentam, para essa amostra, a melhor razão adimplentes/inadimplentes. Os diversos tipos de emprego foram representados por um número inteiro, entre o intervalo $[1,3]$, conforme o seu poder discriminante. Os valores são apresentados na tabela 3.9:

Tabela 3.9: Classificação numérica dos tipos de emprego

Tipo de emprego	Classificação numérica
Emprego privado	1
Sem vínculo empregatício	2
Emprego público	3

Os valores utilizados por essa representação numérica também foram normalizados, de modo a que passassem a se encontrar dentro do intervalo $[0,1]$. Para que esse objetivo fosse alcançado, dividiu-se o valor de cada representação numérica por 3, chegando-se aos valores apresentados na tabela 3.10:

Tabela 3.10: Classificação numérica normalizada dos tipos de emprego

Tipo de emprego	Classificação numérica
Emprego privado	0.33
Sem vínculo empregatício	0.66
Emprego público	1.00

Pensou-se em outra possibilidade para a classificação numérica de dados não métricos; esta outra forma de atribuir valores numéricos aos dados qualitativos tipo de emprego e estado civil também levou em conta a razão adimplentes/inadimplentes apresentada por estas características dos tomadores de crédito das amostras de treinamento, e todo o processo se traduziu em normalizar os valores apresentados para a razão adimplentes/inadimplentes, de modo que os valores a serem utilizados passassem a se encontrar dentro do intervalo $[0,1]$.

De acordo com os valores apresentados na tabela 3.1, o maior valor apresentado para a razão adimplentes/inadimplentes no tocante ao estado civil é 4, de modo que todos os valores apresentados para esta razão foram normalizados, por meio da divisão por 4, passando a ter os valores apresentados na tabela 3.11:

Tabela 3.11: Segunda classificação numérica normalizada dos estados civis

Estado civil	Razão adimplentes/inadimplentes	Razão adimplentes/inadimplentes (normalizada)
Separados	0,636	0,159
Solteiros	0,762	0,1905
Casados	1,188	0,297
Viúvos	4,000	1

A característica tipo de emprego também teve uma representação numérica obtida de forma análoga ao exposto no parágrafo anterior. Conforme exposto na tabela 3.2, o valor mais levado para a razão adimplentes/inadimplentes para esta característica foi apresentado para o os tomadores de crédito que possuíam emprego público, apresentando valor 2,933. Então, todos os valores foram normalizados para o intervalo [0,1], pela divisão por 2,933, conforme apresentamos na tabela 3.12:

Tabela 3.12: Segunda classificação numérica normalizada dos tipos de emprego

Tipo de emprego	Razão adimplentes/inadimplentes	Razão adimplentes/inadimplentes (normalizada)
Privado	0,636	0,159
S/ vínculo de emprego	0,762	0,1905
Público	1,188	0,297

A classificação numérica em questão foi utilizada para o treinamento de uma rede neural supervisionada, mas não foi atingida a convergência, de modo que, para o treinamento e teste das redes neurais utilizadas neste trabalho, foram utilizados os valores de representação numérica para as características tipo de emprego e estado civil apresentados nas tabelas 3.8 e 3.10.

Como o atributo renda recebida mensalmente pelos clientes da carteira que forneceu a amostra em questão possuía seu valor máximo de 5.000,00, o dado numérico salário atual foi normalizado para o intervalo [0, 1], empregando-se a fórmula:

$$Y = (X/5000),$$

Onde:

Y = valor normalizado de X

X = valor efetivo de cada variável.

O objetivo da normalização é diminuir a influência causada por valores que se destacam excessivamente em relação aos demais, ou seja, diminuir a distância entre os valores de variáveis muito espaçadas.

A verificação do grau de associação entre os três possíveis pares de atributos foi efetuada através do cálculo da correlação e da covariância, conforme disposto a seguir na tabela 3.8.

Tabela 3.13: correlação linear entre os pares de atributos

Par de atributos	Correlação linear	Covariância
Renda/ estado civil	0,041338	0,002742
Renda/ tipo de emprego	0,217525	0,011346
Tipo de emprego/ estado civil	0,344303	0,02431

3.4 Implementação dos modelos de Redes Neurais Artificiais

3.4.1 Rede Neural Supervisionada

3.4.1.1 Arquitetura

Após os atributos terem sido pré-processados conforme descrito no item 3.3.2, foi iniciada a construção de uma rede neural supervisionada (*feedforward*), com algoritmo de treinamento *backpropagation*, no qual os pesos são ajustados baseados na regra de aprendizagem por correção de erro. A rede, modelada e executada no código computacional MATLAB[®] com o *toolbox* de redes neurais, foi estruturada em três camadas, ou seja, apenas uma camada oculta foi utilizada. Segundo Haykin (2001), a presença de uma ou mais camadas ocultas habilita a rede a extrair estatísticas de ordem elevada.

Foram testadas diferentes arquiteturas, variando-se os parâmetros internos da rede, visando, desse modo, alcançar uma performance eficiente. A arquitetura considerada mais eficiente será explanada a seguir:

A camada de entrada é formada por três neurônios, um para cada variável de entrada da rede. O número de neurônios na segunda camada foi sendo modificado ao longo dos testes, chegando à convergência com cinco neurônios nesta camada. Na terceira camada há apenas um neurônio que representa a variável de saída da rede: o *score* (valor de saída) obtido. Além do número de neurônios da camada oculta, a taxa de aprendizagem, e o número de épocas também foram sendo modificados durante a realização de inúmeros testes na busca da combinação que produzisse o melhor resultado.

A soma esperada dos quadrados dos erros entre os valores obtidos e os desejados é igual a 0,001.

Como critério de parada utilizou-se o número de épocas igual a 10.000, ou seja, todo o conjunto de treinamento foi apresentado no máximo 10.000 vezes para que a rede convirja para o erro estabelecido sob uma taxa determinada em 0,01.

A função sigmóide (sig) foi definida como a função de ativação para os neurônios das três camadas.

É importante ressaltar que a combinação que resulta na melhor previsão para uma ação não é a mesma combinação que gera o melhor resultado para outra. Portanto, é necessário configurar uma rede neural com parâmetros diferentes para cada ação.

3.4.1.2 Treinamento da rede supervisionada

Segundo Haykin (2001, p.75), “A propriedade que é de importância primordial de uma rede neural é a sua habilidade de aprender”. Uma RNA deve ser treinada para que o objetivo de aprender seja alcançado.

Segundo HAYKIN (2001), os pesos dos neurônios artificiais são os responsáveis pelo armazenamento do conhecimento nas RNA, e devem, portanto, ser ajustados por meio do processo de treinamento, de modo a tornar a rede apta a fornecer respostas de forma mais correta possível, quando novas informações lhe forem apresentados em uma etapa posterior à etapa de treinamento, chamada de etapa de teste.

Na fase de treinamento são apresentados padrões à rede. A rede tenta assimilar as características desses padrões, aprendendo os processos neles contidos. Cada elemento da amostra de treinamento e da amostra de teste é apresentado à rede

como um vetor de três atributos: renda mensal, tipo de emprego e estado civil. O conjunto de dados de uma das amostras passou a ter o formato apresentado na tabela 3.12:

Tabela 3.14: exemplo de dados de uma amostra para treinamento

Renda normalizada:	Tipo de emprego:..	Estado civil:
0,3102	0,33	0,50
0,5920	0,33	0,50
0,0610	0,33	0,75

As amostras usadas para o treinamento da rede neural consistem em um total de 100 tomadores de crédito (50 com histórico de adimplência e 50 com histórico de inadimplência) enquanto outros 30 tomadores de crédito (15 adimplentes e 15 inadimplentes) foram usados para a etapa de testes. Cabe ressaltar que os elementos para a amostra de testes foram escolhidos aleatoriamente e foi mantida a mesma proporção de classes presente na amostra de treinamento (50% de adimplentes e 50% de inadimplentes).

Uma vez definidos os padrões de entrada, chegou-se então à etapa de decisão do valor dos “targets”, ou seja, o valor final que cada neurônio da camada de saída deverá apresentar como resultado. Iniciou-se o treinamento fixando os valores dos “targets” em 0,1 para os cinquenta candidatos conhecidamente inadimplentes e com o valor 0,9 para os sabidamente adimplentes. Não houve convergência. De modo a tentar conseguir com que rede alcançasse a convergência, a distância entre os valores esperados de saída da rede para os adimplentes e para os inadimplentes foi sendo gradativamente diminuída, e o treinamento foi novamente tentado, testando-se diferentes valores para os “targets”, até que fossem encontrados valores tais que

permitissem ocorrer a convergência. Foram testados os seguintes pares de valores de saída:

Tabela 3.15: valores esperados de saída testados

	Inadimplentes	Adimplentes	Convergência
Valor esperado de saída	0,20	0,80	Não
Valor esperado de saída	0,30	0,70	Não
Valor esperado de saída	0,40	0,60	Não
Valor esperado de saída	0,45	0,55	Sim

Os valores esperados de saída com os quais a rede alcançou a convergência foram 0,45 e 0,55, para os inadimplentes e para os adimplentes, respectivamente. Prosseguiu-se então com o treinamento da RNA.

Quanto ao número de neurônios, como os padrões de entrada possuem três atributos, a rede possui esse mesmo número de neurônios na primeira camada (camada de entrada), e inicia-se o treinamento com um neurônio na camada intermediária, e tenta-se a convergência. Na medida em que a convergência não foi conseguida, procedendo-se à alteração do número de neurônios na camada intermediária até que chegou-se à convergência. Este objetivo foi alcançado com cinco neurônios na camada escondida. A partir de então, cada uma das quatro amostras foi apresentada à rede neural até que a convergência fosse alcançada 33 vezes, obtendo então 33 conjuntos diferentes de pesos,. Como as amostras utilizadas foram apresentadas à rede até que a convergência fosse alcançada 33 vezes, pode-se admitir, baseado no Teorema do Limite Central, que os resultados encontrados distribuem-se em torno de um valor médio, e que é possível o emprego de uma distribuição normal padronizada para a realização de análises estatísticas.

3.4.1.3 Teste da rede supervisionada

Realizado o treinamento, o próximo passo foi efetuar o teste das redes. O teste (ou verificação) é a fase onde atributos de elementos que não pertencem à amostra de treinamento são apresentados à rede, para verificar o desempenho do aprendizado desta.

Os atributos de quatro grupos distintos trinta tomadores de crédito foram apresentados à rede, para que fosse efetuada a classificação individual de cada um dos membros da amostra de teste como provável adimplente ou provável inadimplente. Como já era conhecido *a priori*, quinze candidatos se tornaram inadimplentes em suas operações de crédito, e os outros quinze mantiveram-se adimplentes, de modo que esse era o número a ser comparado após a etapa de teste. Como já foi explanado anteriormente, os inadimplentes deveriam apresentar o *score* (valor apresentado pelo neurônio da camada de saída) 0,45, enquanto era esperado que os adimplentes apresentassem como *score* o valor de 0,55. Uma vez atingida a convergência, foram calculadas as médias e os desvios padrões dos valores de saída obtidos pelos candidatos inadimplentes e adimplentes utilizados nas quatro amostras utilizadas para treinamento. As médias dos valores encontrados são abaixo apresentadas:

Tabela 3.16: valores de média e desvio padrão para os inadimplentes (amostra 1)

Amostra 1	Inadimplentes
Média	0,4714
Desvio padrão	0,019116
Média + desvio padrão	0,4905

Tabela 3.17: valores de média e desvio padrão para os adimplentes (amostra 1)

Amostra 1	Adimplentes
Média	0,5280
Desvio padrão	0,023188
Média - desvio padrão	0,5048

Tabela 3.18: valores de média e desvio padrão para os inadimplentes (amostra 2)

Amostra 2	Inadimplentes
Média	0,4704
Desvio padrão	0,027993
Média + desvio padrão	0,4984

Tabela 3.19: valores de média e desvio padrão para os adimplentes (amostra 2)

Amostra 2	Adimplentes
Média	0,5298
Desvio padrão	0,019777
Média - desvio padrão	0,5101

Tabela 3.20: valores de média e desvio padrão para os inadimplentes (amostra 3)

Amostra 3	Inadimplentes
Média	0,4694
Desvio padrão	0,0090
Média + desvio padrão	0,4784

Tabela 3.21: valores de média e desvio padrão para os adimplentes (amostra 3)

Amostra 3	Adimplentes
Média	0,5290
Desvio padrão	0,032567
Média - desvio padrão	0,4965

Tabela 3.22: valores de média e desvio padrão para os inadimplentes (amostra 4)

Amostra 4	Inadimplentes
Média	0,4691
Desvio padrão	0,010478
Média + desvio padrão	0,4795

Tabela 3.23: valores de média e desvio padrão para os adimplentes (amostra 4)

Amostra 4	Adimplentes
Média	0,5231
Desvio padrão	0,027691
Média - desvio padrão	0,4954

Tabela 3.24: valores médios de média e desvio padrão para os inadimplentes (das quatro amostras)

Média das amostras	Inadimplentes
Média	0,470075
Desvio padrão	0,018302
Média + desvio padrão	0,488377

Tabela 3.25: valores médios de média e desvio padrão para os adimplentes (das quatro amostras)

Média das amostras	Adimplentes
Média	0,527475
Desvio padrão	0,026250
Média - desvio padrão	0,501225

Após os cálculos dos valores médios das médias e dos desvios padrões, estabeleceu-se que os candidatos a serem classificados como prováveis inadimplentes deveriam apresentar o valor de saída menores que 0,488377, que é o valor médio das médias dos resultados de saída obtido pelos inadimplentes somados a seu desvio padrão. Determinou-se, também, que os candidatos a serem classificados como adimplentes deveriam apresentar os valores de saída maiores que 0,501225, que foi obtido pela soma do valor médio dos resultados de saída obtidos pelos adimplentes, subtraído de seu desvio padrão.

O teste do χ^2 foi utilizado comparar as divergências entre as frequências observadas e as esperadas.

Uma vez calculado o χ^2 médio para cada amostra, procurou-se na tabela de distribuição de χ^2 o valor do χ^2 crítico considerando o nível de significância adotado e os graus de liberdade.

Quando analisados os resultados apresentados pelo teste das quatro amostras pela rede neural supervisionada foram apresentados os resultados médios de χ^2 e de percentual de acertos:

Tabela 3.26: Média de acertos e de χ^2 para a Amostra 1

χ^2	0,922
Desvio padrão do χ^2	0,271
Média de acertos(%)	83,61
Desvio padrão da média de acertos	2,585

Tabela 3.27: Média de acertos e de χ^2 para a Amostra 2

χ^2	1,648
Desvio padrão do χ^2	0,480
Média de acertos(%)	80,00
Desvio padrão da média de acertos	2,615

Tabela 3.28: Média de acertos e de χ^2 para a Amostra 3

χ^2	3,768
Desvio padrão do χ^2	0,904
Média de acertos(%)	74,65
Desvio padrão da média de acertos	3,327

Tabela 3.29: Média de acertos e de χ^2 para a Amostra 4

χ^2	0,730
Desvio padrão do χ^2	0,581
Média de acertos(%)	86,97
Desvio padrão da média de acertos	4,741

Para $v = 1$, $\chi^2_{0,95} = 3,84$. Como o χ^2 obtido foi menor que o χ^2 crítico, então, visto que $0,5333 < 3,84$, não se pode rejeitar a hipótese de que o modelo para previsão de inadimplência seja eficiente, para o nível de significância de 0,05.

3.4.2 Rede Neural Não Supervisionada

Nesta seção apresenta-se uma descrição de um método de previsão de inadimplência desenvolvido para as redes auto-organizáveis de Kohonen. Os métodos aqui apresentados diferem daqueles descritos e desenvolvidos nos capítulos anteriores por constituírem técnicas de treinamento não-supervisionado.

3.4.2.1 PRÉ-PROCESSAMENTO DOS DADOS

Todo o pré-processamento de dados descrito no item 3.3.2, página 33 se aplica a essa etapa, ou seja, os dados pré-processados de quatro amostras de 100 tomadores de crédito utilizados para as amostras de treinamento e de 30 clientes com operações creditícias apresentados para o teste da rede neural supervisionada também foram utilizados para o treinamento e teste da rede Kohonen.

3.4.2.2 Arquitetura

Número de neurônios:

A rede apresenta apenas uma linha, com duas colunas, tendo portanto dois neurônios, onde cada um representa uma classificação dos candidatos a crédito, e a rede realizará os agrupamentos de acordo com as similaridades dos padrões de

entrada. Foi então estabelecido que a rede agrupasse os elementos da amostra de treinamento em dois grupos: prováveis adimplentes e prováveis inadimplentes.

Definição da vizinhança:

O próximo passo foi determinar o raio da vizinhança, variável que tem fundamental importância no processo de atualização/correção dos pesos. Devido ao número relativamente pequeno de neurônios, ficou determinado que a vizinhança seria toda a rede, ou seja, o neurônio vencedor seria o que sofreria a maior correção de seu peso, enquanto que o outro neurônio teria também seus pesos corrigidos, de forma inversamente proporcional à sua distância do neurônio vencedor.

Parâmetros para correção do neurônio vencedor e de seus vizinhos:

O neurônio vencedor e seus vizinhos foram corrigidos seguindo os parâmetros de vizinhança previamente determinados. Dois parâmetros de extrema importância nesta etapa foram as variáveis α e σ . Ambas estão diretamente ligadas com o montante de conhecimento que cada neurônio aprende em relação ao neurônio vencedor.

A variável α , também chamada de taxa de aprendizagem, teve seu valor inicial em 0,9. Ao final de cada iteração, o valor da taxa de aprendizagem foi corrigido, da seguinte forma:

- Enquanto a taxa de aprendizagem fosse maior que 0,05 (relativamente grande), teria seu valor alterado para 95% do valor original;

- Caso o valor da taxa de aprendizagem já se encontrasse igual ou menor que 0,05 (relativamente pequeno), teria seu valor alterado para 10% do valor original, acelerando a convergência.

A variável σ teve seu valor fixado em 3.

Número de épocas:

Cada época do treinamento corresponde à apresentação de todos os vetores de entrada para a rede. O critério de parada utilizado foi o número máximo de épocas de 15.000.

Observação: As variáveis σ , α , e o número de épocas não possuem valores pré-determinados a serem utilizados. Os valores utilizados poderiam ser alterados durante as experimentações, caso os resultados apresentados não fossem satisfatórios.

3.4.2.3 Treinamento da rede não supervisionada

Uma vez definida a arquitetura da rede não supervisionada, iniciou-se o treinamento da RNA, seguindo o algoritmo e equações do item 2.3 (capítulo 2). As quatro amostras de treinamento foram apresentadas à rede, e após as 15.000 iterações estabelecidas como critério de parada, a rede nos apresentou os valores de cada um dos três pesos dos dois neurônios, devidamente corrigidos. Esses valores são apresentados nas tabelas a seguir:

Tabela 3.30: Pesos obtidos após o treinamento com a amostra 1

	Elemento 1	Elemento 2	Elemento 3
Neurônio 1	0,405916	0,390510	0,525758
Neurônio 2	0,429795	0,926906	0,740042

Tabela 3.31: Pesos obtidos após o treinamento com a amostra 2

	Elemento 1	Elemento 2	Elemento 3
Neurônio 1	0,378123	0,458343	0,564002
Neurônio 2	0,432635	0,958688	0,737840

Tabela 3.32: Pesos obtidos após o treinamento com a amostra 3

	Elemento 1	Elemento 2	Elemento 3
Neurônio 1	0,259963	0,927325	0,649363
Neurônio 2	0,182402	0,380830	0,547466

Tabela 3.33: Pesos obtidos após o treinamento com a amostra 4

	Elemento 1	Elemento 2	Elemento 3
Neurônio 1	0,378123	0,452709	0,572475
Neurônio 2	0,432635	0,958688	0,725720

3.4.2.4 Teste da rede não supervisionada

Concluído o treinamento, o próximo passo foi realizar o teste da rede. O teste (ou verificação) é a fase onde atributos de elementos que não pertencem à amostra de treinamento são apresentados à rede, para verificar o desempenho do aprendizado desta.

Quatro amostras de teste, cada uma com trinta elementos foram utilizadas. Os atributos dos trinta candidatos a crédito foram apresentados à rede, para que fossem calculadas as distâncias entre os atributos dos candidatos a crédito e os elementos dos vetores peso. Houve a “clusterização” (agrupamento), e os valores de percentual de acertos e os valores calculados de χ^2 são apresentados nas tabelas abaixo:

Tabela 3.34 Percentual de acertos e valor de χ^2 para a amostra 1 (treinamento não supervisionado)

χ^2	0,133
Acertos(%)	93,33

Tabela 3.35: Percentual de acertos e valor de χ^2 para a amostra 2 (treinamento não supervisionado)

χ^2	0,270
Média de acertos(%)	93,33

Tabela 3.36: Percentual de acertos e valor de χ^2 para a amostra 3 (treinamento não supervisionado)

χ^2	0,530
Média de acertos(%)	86,67

Tabela 3.37: Percentual de acertos e valor de χ^2 para a amostra 4 (treinamento não supervisionado)

χ^2	0,530
Média de acertos(%)	86,67

Para $v = 1$, $\chi^2_{0,95} = 3,84$. Como o valor do χ^2 obtido foi menor que o χ^2 crítico para todas as amostras, então não se pode rejeitar a hipótese de que o modelo para previsão de inadimplência seja eficiente, para o nível de significância de 0,05.

De posse destas informações, chegou-se à conclusão que um dos dois neurônios da rede tem forte incidência de ativação quando os atributos de um candidato com características ligadas à inadimplência são expostos à rede, enquanto que o outro neurônio teria sua ativação intimamente ligada à exposição de atributos de um candidato com histórico de adimplência.

3.5. RESULTADOS

Serão expostos abaixo quadros comparativos com os percentuais médios de acerto e valores médios de χ^2 para cada uma das quatro amostras, quando tratadas pela rede supervisionada, e percentuais de acertos e valores de χ^2 , quando a rede foi tratada pela rede não supervisionada.

Tabela 3.38: Comparativo de percentual de acertos e valor de χ^2 para a amostra 1

	Rede não supervisionada:	Rede supervisionada:
χ^2	0,133	0,922
Média de acertos (%)	93,33	83,61

Tabela 3.39: Comparativo de percentual de acertos e valor de χ^2 para a amostra 2

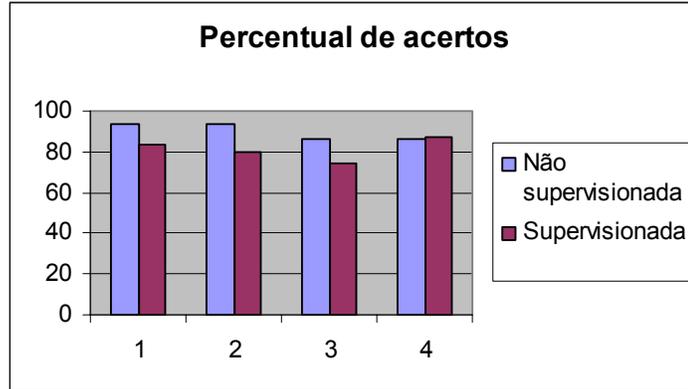
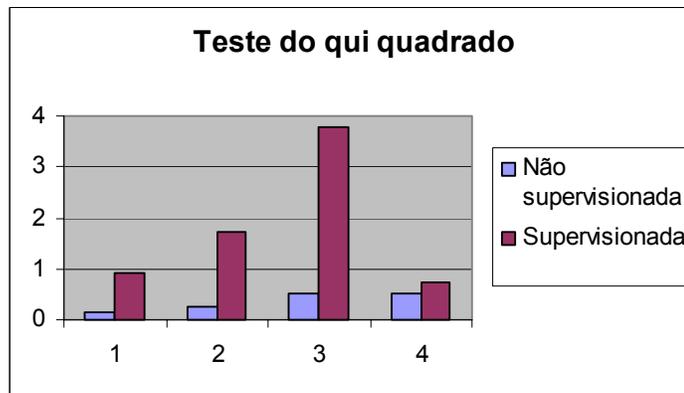
	Rede não supervisionada:	Rede supervisionada:
χ^2	0,270	1,723
Média de acertos (%)	93,33	79,48

Tabela 3.40: Comparativo de percentual de acertos e valor de χ^2 para a amostra 3

	Rede não supervisionada:	Rede supervisionada:
χ^2	0,530	3,768
Média de acertos(%)	86,67	74,77

Tabela 3.41: Comparativo de percentual de acertos e valor de χ^2 para a amostra 4

	Rede não supervisionada:	Rede supervisionada:
χ^2	0,530	0,73
Média de acertos(%)	86,67	86,99

Figura 3.1: Gráfico comparativo de percentual de acertos para as quatro amostras**Figura 3.2:** Gráfico comparativo de valores de χ^2 para as quatro amostras

Como o valor crítico de χ^2 para o nível de significância 0,05 é de 3,84, podemos concluir que os modelos implementados têm validade estatística, pois mesmo a amostra que apresentou os piores resultados (amostra 3, quando tratada pela rede supervisionada) apresentou um valor de inferior ao valor crítico do χ^2 .

CAPÍTULO 4

CONCLUSÕES E TRABALHOS FUTUROS

Os resultados obtidos demonstram que tanto as redes neurais supervisionadas quanto as redes neurais não supervisionadas mostraram-se instrumentos eficazes para o processo de previsão de inadimplência. Ao testarmos as redes com as quatro amostras, foram conseguidos índices de acerto que, quando submetidos ao teste do qui-quadrado, tornaram positiva a validação estatística dos modelos. É também informação relevante o fato de que em todas as amostras a rede não supervisionada atingiu desempenho superior ao alcançado pela rede supervisionada, conforme pode-se observar por meio da avaliação dos resultados apresentados para o teste do qui-quadrado, e também dos valores apresentados no comparativo dos percentuais médios de acerto, vide tabelas 3.34, 3.35, 3.36 e 3.37.

As redes neurais implementadas revelaram-se uma ferramenta muito útil para auxílio à gestão de crédito financeiro, permitindo que se realizem previsões sobre as expectativas de retornos dos valores envolvidos nas operações de crédito com uma margem de confiabilidade nos resultados.

Foi possível, por meio da utilização das redes neurais, no conjunto de informações utilizadas, classificar corretamente até 93,33% dos candidatos a crédito (resultado conseguido com o uso da rede não supervisionada). As redes neurais supervisionadas conseguiram classificar corretamente até 86,99% dos candidatos a crédito.

As redes neurais não supervisionadas apresentaram maior eficiência do que as redes supervisionadas, indicando seu uso como ferramenta para a classificação de candidatos a crédito como prováveis adimplentes ou inadimplentes. Além da maior eficiência, também cabe ressaltar a maior regularidade da rede não supervisionada, que, sendo treinada e testada com utilização de quatro amostras diferentes (assim como a rede supervisionada), apresentou média de acertos variando de 86,67% a 93,33%, enquanto que a rede supervisionada apresentou média de acertos variando de 74,77% a 86,99%. De forma similar, as redes não supervisionadas também apresentaram menor variação de valores apresentados como resultado como teste do χ^2 , variando de 0,133 a 0,570, enquanto que a rede supervisionada apresentou resultados de χ^2 variando de 0,730 a 3,768. Cabe salientar que uma classificação errônea feita pela rede poderia ter como consequência a negação de crédito a um cliente que seria um provável adimplente, ou a concessão de crédito a um candidato provável inadimplente, onde, em ambas as hipóteses, a consequência seria a diminuição do lucro da instituição financeira. A confiabilidade e precisão da rede utilizada são fundamentais no processo decisório, e, então, a rede não supervisionada, cabe enfatizar, apresentou-se como uma ferramenta mais eficiente do que a rede supervisionada, como pode-se concluir após a comparação dos resultados obtidos para a o teste do χ^2 e dos valores de média de percentual de acertos, já informados neste parágrafo.

Os resultados obtidos com as redes neurais não supervisionadas e com as redes neurais supervisionadas desenvolvidas, demonstram que estes tipos de rede neural conseguem reconhecer padrões existentes nas características de candidatos a crédito possibilitando, então, a classificação já citada no parágrafo anterior. Isto

confirma a idéia de que as redes neurais podem funcionar como instrumentos de previsão de inadimplência.

É interessante notar que, mesmo fazendo-se uso de somente três atributos, consegue-se sucesso no treinamento e teste de ambas as redes, fator decisivo para que as redes desenvolvessem a capacidade de classificação. A escolha de variáveis com alto poder discriminante apresentou-se eficiente para passar às redes informações relevantes sobre os candidatos a crédito, e, ao mesmo tempo, lidar com um número reduzido de informações.

TRABALHOS FUTUROS

Estudos futuros podem ser realizados, procurando identificar outras formas de construir as redes. A pesquisa envolverá também a escolha de outras variáveis que permitam distinguir de maneira mais precisa candidatos adimplentes e inadimplentes. A utilização da relação renda recebida e renda comprometida mensalmente poderá ser mais eficiente do que a variável renda isoladamente.

Pode ser adicionado, futuramente, um estudo da implementação do mesmo tipo de rede utilizada neste trabalho, formando, porém, um maior número de classes de clientes, principalmente no tocante à renda. Poder-se-ia “estretitar o foco”, tentando lidar cada vez mais com candidatos mais parecidos, e que, ainda assim, apresentam resultados opostos no tocante à adimplência. O treinamento da rede com, por exemplo, tomadores de crédito com a mesma faixa de renda, com demais características diferentes, e históricos distintos de adimplência/inadimplência talvez leve a RNA a detectar padrões ligados a inadimplência em tomadores de crédito com perfis muito semelhantes.

O desenvolvimento de uma rede auto-organizável do tipo ART-2 (*Adaptive Resonance Theory*) também é uma possibilidade de trabalho futuro. A rede ART se caracteriza principalmente pela sua capacidade em manter em equilíbrio as propriedades de plasticidade (discriminação) e estabilidade (generalização). A rede ART seria capaz de, através de seu treinamento com as amostras de candidatos a crédito, informar o número de classes para o estudo em questão. Talvez possa-se trabalhar com classes como “alta tendência à inadimplência”, “média tendência à inadimplência”, “média tendência à adimplência” e “alta tendência à adimplência”.

O uso de um sistema híbrido associando lógica fuzzy às redes neurais é uma outra possibilidade para prorrogação deste estudo, devidos à característica *fuzzy* dos valores de saída obtidos como resultado das redes supervisionadas.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ALLEN, L. Credit risk modelling of middle markets, *Zicklin School of Business*, Baruch College, CUNY, 2002.
- [2] ALTMAN, E. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, v.23, n. 4, p. 589-609, 1968.
- [3] ALTMAN, E. I., HALDEMAN, R. Corporate credit scoring models: approaches and tests for a successful implementation, Journal of Commercial Lending, p. 10-22, May 1995.
- [4] APTÉ, C., KASTNER, J. Financial Applications. Special Issue, *IEEE Expert*, vol. 2, n. 3, 1987.
- [5] BACK, Barbro et al, Choosing bankruptcy predictors using discriminant analysis, logit analysis and genetic algorithms, *Turku Centre for Computer Science Technical report N° 40, ISBN 951-650-828-6, ISSN 1239-1891*, 1996.
- [6] BARTH, N. L. Inadimplência: construção de modelos de previsão, São Paulo: Nobel, 2004
- [7] BRAGA, A.P., CARVALHO, A.C.P.L., LUDERMIR, T.B. Redes Neurais Artificiais: Teoria e aplicações, Rio de Janeiro: LTC, 2000.
- [8] BRIGHAM, E. GAPENSKI, L. EHRHARDT, M. Administração Financeira. São Paulo: Atlas, 2001
- [9] CARTER, C., CATLETT, J. Assessing Credit Card Applications Using Machine Learning, *IEEE Expert*, vol. 2, n. 3, 1987.
- [10] FAUSETT, L. Fundamentals of neural networks: architectures, algorithms, and applications. Upper Saddle River, NJ: Prentice Hall, 1994.

- [11] GUJARATI, D.N. Basic econometrics. 3rd. Ed. Singapore: McGraw-Hill, 1995.
- [12] HAYKIN, S. Redes Neurais: princípios e prática. Tradução Paulo Martins Engel. 2 ed. Porto Alegre: Bookman, 2001.
- [13] KOHONEN, T. Self-Organized Formation of Topologically Correct Feature Maps, Biological Cybernetics, 1982.
- [14] KOVÁCS, Z. L. Redes Neurais Artificiais. 2ª Edição. São Paulo: Collegium, Cognito, 1996.
- [15] KOVACS, Z. L. O Cérebro e a sua Mente: Uma introdução à neurociência computacional. São Paulo: Edições Acadêmicas, 1997.
- [16] MAGALHÃES, M.N., LIMA, A.C.P. Noções de probabilidade e estatística. São Paulo: da Universidade de São Paulo: 2002.
- [17] MATARAZZO, D. C. Análise Financeira de Balanços; São Paulo: Atlas, 1995.
- [18] Neural Network Toolbox User's Guide. (2000). Math Works Inc.
- [19] SAUNDERS, A. Métodos tradicionais e novos de gerenciamento de risco de crédito, IV congresso nacional de crédito, São Paulo: 1996.
- [20] SCHIRICKEL, W. Análise de crédito: concessão e gerência de empréstimos. 3ª ed. São Paulo: Atlas, 1997.
- [21] SECURATO, J. Decisões financeiras em condições de risco, São Paulo: Atlas, 1996.
- [22] SILVA, J. Análise e decisão de crédito, São Paulo, Atlas: 1988.
- [23] SPIEGEL, M. R. Estatística: resumo da teoria. São Paulo: McGraw-Hill do Brasil, 1977.

- [24] STEINER, M. T. A.; CARNIERI, C.; STEINER NETO, P. J.; KOPITTKE, B. Sistemas especialistas probabilístico e redes neurais na análise do crédito bancário. *Revista de Administração*, São Paulo, v. 34, n. 3, p. 56-67, jul./set. 1999.
- [25] TAFNER, M. , XEREZ, M, RODRIGUES FILHO, I. Redes Neurais Artificiais, Introdução e princípios de neurocomputação. Blumenau: Furb, 1995.
- [26] THOMPSON, C.A., CAETANO, F.M., THOMPSON M.F. Redes neurais artificiais e padrões de falência, Revista de Economia e Administração, Rio de Janeiro, volume 3, n. 3, p. 225-242, 2004.
- [27] TRELEAVEN, P., GONNATILAKE, S. Intelligent Financial Technologies, *Proc. of the Workshop on Parallel Problem Solving From Nature: Applications in Statistics and Economics*,1992.
- [28] TRIPPI, R. TURBAN, E. Neural Networks in Finance and Investing . Chicago: Probus Publishing Company, 1993.
- [29] REED, E. W.; GILL, E. K. Bancos Comerciais. São Paulo: Makron Books, 1994.
- [30] WASSERMAN, P. D. Neural Computing: theory and practice. New York: Van Nostrand Reinhold, 1989.
- [31] WASSERMAN, P. D. Advanced Methods in Neural Computing. New York: Van Nostrand Reinhold, 1993.

APÊNDICE

A: ALGORITMO DE RETROPROPAGAÇÃO (*BACKPROPAGATION*)

O sinal de erro na saída do neurônio j , na iteração n , é definido por:

$$e_j(n) = t_j(n) - o_j(n), \text{ o neurônio } j \text{ é um nó de saída.} \quad (\text{A.1})$$

Definimos o valor instantâneo da energia de erro para o neurônio j como: $\frac{1}{2}e_j^2(n)$.

O valor instantâneo $E(n)$ da energia total do erro é obtido somando-se todos os termos $\frac{1}{2}e_j^2(n)$ de todos os neurônios da camada de saída da rede.

$$E(n) = \frac{1}{2} \sum_{j \in C} e_j^2(n) \quad (\text{A.2})$$

C inclui todos os neurônios da camada de saída.

A energia média do erro quadrado é obtida somando-se todos os $E(n)$ para todos os n e então normalizando em relação ao tamanho do conjunto N , conforme abaixo:

$$E_{med} = \frac{1}{N} \sum_{n=1}^N E(n) \quad (\text{A.3})$$

Um neurônio j é alimentado por um conjunto de sinais funcionais produzidos por uma camada de neurônios à sua esquerda. O campo local induzido $v_j(n)$ produzido na entrada da função de ativação associada ao neurônio j é portanto:

$$v_j(n) = \sum_{i=0}^m w_{ji}(n) y_i(n), \quad (\text{A.4})$$

onde m é o número total de entradas aplicadas ao neurônio j . Assim, o sinal funcional $o_j(n)$ que aparece na saída do neurônio j na iteração n é

$$o_j(n) = \varphi_j(v_j(n)) \quad (\text{A.5})$$

O algoritmo de retropropagação aplica uma correção $\Delta w_{ji}(n)$ ao peso sináptico

$w_{ji}(n)$, que é proporcional à derivada parcial $\frac{\partial E(n)}{\partial w_{ji}(n)}$. De acordo com a regra da

cadeia do cálculo tem-se:

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial o_j(n)} \frac{\partial o_j(n)}{\partial v_j(n)} \frac{\partial v_j(n)}{\partial w_{ji}(n)} \quad (\text{A.6})$$

Diferenciando-se ambos os lados da equação (A.2) em relação a $e_j(n)$, obtém-se

$$\frac{\partial e_j(n)}{\partial o_j(n)} = e_j(n) \quad (\text{A.7})$$

Diferenciando-se ambos os lados da equação (A.1) em relação a $o_j(n)$, obtém-se

$$\frac{\partial e_j(n)}{\partial o_j(n)} = -1 \quad (\text{A.8})$$

A seguir, diferenciando-se ambos os lados da equação (A.5) em relação a $v_j(n)$,

obtém-se

$$\frac{\partial o_j(n)}{\partial v_j(n)} = \varphi'_j(v_j(n)) \quad (\text{A.9})$$

Finalmente, diferenciar a equação (A.4) em relação a $w_{ji}(n)$ produz

$$\frac{\partial v_j(n)}{\partial w_{ji}(n)} = o_i(n) \quad (\text{A.10})$$

O uso das equações de (A.7) a (A.10) e, (A.6) produz

$$\frac{\partial E(n)}{\partial w_{ji}(n)} = e_j(n) \varphi'_j(v_j(n)) o_i(n) \quad (\text{A.11})$$

A correção $\Delta w_{ji}(n)$ aplicada a $w_{ji}(n)$ é definida por:

$$\Delta w_{ji} = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} \quad (\text{A.12})$$

Onde η é o parâmetro da taxa de aprendizagem do algoritmo da retropropagação. O uso do sinal negativo na equação (A.12) leva à redução do valor de $E(n)$. O uso da equação (A.11) em (A.12) produz

$$\Delta w_{ji} = -\eta \delta_i(n) o_i(n) \quad (\text{A.13})$$

Onde o gradiente local $\delta_i(n)$ é definido por

$$\begin{aligned} \delta_i(n) &= \frac{\partial E(n)}{\partial v_j(n)} \\ &= \frac{\partial E(n)}{\partial e_j(n)} \frac{\partial e_j(n)}{\partial o_j(n)} \frac{\partial o_j(n)}{\partial v_j(n)} \\ &= e_j(n) \varphi'_j(v_j(n)) \end{aligned} \quad (\text{A.14})$$

Das equações (A.13) e (A.14) nota-se que um fator chave para o cálculo do ajuste de peso $\Delta w_{ji}(n)$ é o sinal de erro $e_j(n)$ na saída do neurônio j . Pode-se então identificar dois casos distintos, dependendo da localização do neurônio j na rede. Caso o neurônio j esteja localizado na camada de saída, ele será suprido com um valor desejado de saída (*target*) particular, fazendo com que o cálculo do erro associado seja direto. Diferentemente, se o neurônio j pertencer a uma camada oculta, não é possível calcular de forma direta o valor da correção $\Delta w_{ji}(n)$, exigindo que seja feita a retropropagação do erro através da rede.

Caso 1: o neurônio j pertence à camada de saída:

Quando o neurônio j está localizado na camada de saída da rede, pode-se utilizar a equação (A.1) para calcular o sinal de erro $e_j(n)$ associado a este neurônio, e, então, calcular diretamente o gradiente local $\delta_i(n)$, utilizando a equação (A.14)

Caso 2: o neurônio j pertence a uma camada oculta:

Quando o neurônio j está localizado em uma camada oculta da rede, não existe um valor de saída esperado para aquele neurônio. Então, o sinal de erro para um neurônio oculto deve ser determinado recursivamente, em termos dos sinais de erro de todos os neurônios aos quais o neurônio oculto está diretamente conectado. De acordo com a equação (A.14), podemos redefinir o gradiente local $\delta_i(n)$ para o neurônio oculto j como

$$\delta_i(n) = \frac{\partial E(n)}{\partial o_j(n)} \frac{\partial o_j(n)}{\partial v_j(n)}$$

(A.15)

$$= \frac{\partial E(n)}{\partial o_j(n)} \varphi'_j(v_j(n)), \text{ onde o neurônio } j \text{ é um neurônio oculto.}$$

Para calcular a derivada parcial $\partial E(n)/\partial o_j(n)$ pode-se proceder como segue abaixo:

$$E(n) = \frac{1}{2} \sum_{k \in C} e_k^2(n), \text{ sendo o neurônio } k \text{ um neurônio da camada de saída.} \quad (\text{A.16})$$

Repare que a equação (A.16) é a equação (A.2) com o índice k no lugar do índice j , que se refere ao neurônio oculto, no caso 2. Diferenciando-se a equação (A.16) em relação ao sinal funcional $o_j(n)$, obtém-se

$$\frac{\partial E(n)}{\partial o_j(n)} = \sum_k e_k \frac{\partial e_j(n)}{\partial o_j(n)} \quad (\text{A.17})$$

A seguir, utiliza-se a regra da cadeia para a derivada parcial e reescrevemos a equação (A.17) na forma equivalente $\partial E(n)/\partial y_j(n)$

$$\frac{\partial E(n)}{\partial o_j(n)} = \sum_k e_k \frac{\partial e_k(n)}{\partial v_k(n)} \frac{\partial v_k(n)}{\partial o_j(n)} \quad (\text{A.18})$$

Mas

$$e_k(n) = t_k(n) - o_k(n) \quad (\text{A.19})$$

$= t_k(n) - \varphi_k(v_k(n))$, sendo o neurônio k da camada de saída.

Assim,

$$\frac{\partial e_k(n)}{\partial v_k(n)} = -\varphi'_k(v_k(n)), \quad (\text{A.20})$$

e

$$v_k(n) = \sum_{j=0}^m w_{kj}(n) o_j(n), \quad (\text{A.21})$$

onde m é o número total de entradas aplicadas ao neurônio k .

Diferenciando-se a equação (A.21) em relação a $\partial o_j(n)$ produz-se

$$\frac{\partial v_k(n)}{\partial o_j(n)} = w_{kj}(n) \quad (\text{A.22})$$

Utilizando-se as equações (A.20) e (A.22) em (A.18) obtém-se a derivada parcial desejada:

$$\frac{\partial E(n)}{\partial o_j(n)} = -\sum_k e_k(n) \varphi'_k(v_k(n)) w_{kj}(n) \quad (\text{A.23})$$

$$= -\sum_k \delta_k(n) w_{kj}(n)$$

Na segunda linha da equação (A.23), utilizou-se a definição de gradiente local dada na equação (A.14), com o índice j substituído por k .

Finalmente, utilizando-se a equação (A.23) em (A.15), obtém-se a *fórmula de retropropagação* para o gradiente local $\delta_j(n)$ como descrito:

$$\delta_j(n) = \varphi'_j(v_j(n)) \sum_k \delta_k(n) w_{kj}(n), \text{ onde o neurônio } k \text{ é de uma camada oculta (A.24)}$$

Resumindo:

$$\begin{pmatrix} \text{Correção} \\ \text{de peso} \\ \Delta w_{ji}(n) \end{pmatrix} = \begin{pmatrix} \text{Parâmetro da} \\ \text{taxa de aprendizagem} \\ \eta \end{pmatrix} \begin{pmatrix} \text{Gradiente} \\ \text{local} \\ \delta_j(n) \end{pmatrix} \begin{pmatrix} \text{Sinal de entrada} \\ \text{do neurônio} \\ y_i(n) \end{pmatrix} \quad (\text{A.25})$$

- Se o neurônio j pertence à camada de saída, $\delta_j(n)$ é igual ao produto da derivada $\varphi'_j(v_j(n))$ pelo sinal de erro $e_j(n)$, ambos sendo associados ao neurônio j , conforme a equação (A.14).

- Se o neurônio j pertence a uma camada oculta, $\delta_j(n)$ é igual ao produto da derivada $\varphi'_j(v_j(n))$ associada pela soma ponderada dos δ s calculados na próxima camada oculta ou camada de saída que estiverem conectadas ao neurônio j ; veja a equação (A.24).

B: FUNDAMENTOS DE ESTATÍSTICA

B.1: Medidas de posição ou de tendência central:

B.1.1: Média aritmética

Caso uma variável X apresente n valores observados, representados por x_1, x_2, \dots, x_n , a média aritmética é a soma dos valores observados, dividida pelo número de observações.

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{\sum_{i=1}^n x_i}{n} \quad (\text{B.1})$$

B.2: Medidas de dispersão

Quando desejamos quantificar a variabilidade dos valores de uma variável, fazemos uso das medidas de dispersão.

B.2.1: Variância:

É calculada pelo quadrado médio dos desvios em relação à média. A variância, referente à variável X de um conjunto de dados, é definida por:

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} \quad (\text{B.2})$$

B.2.2: Desvio padrão :

É a raiz quadrada da variância , ou seja

$$\sigma = \sqrt{\sigma^2} \quad (\text{B.3})$$

B.3: Medidas de associação:

As medidas de associação buscam explicar como se comportam umas variáveis em função de outras.

B.3.1: Covariância:

É o valor médio do produto dos desvios de X e Y , em relação às suas respectivas médias. Utiliza-se a covariância para determinar a relação entre dois conjuntos de dados, indicando a tendência das variáveis de moverem-se na mesma direção, isto é, indicando se uma variável cresce caso uma outra também cresça.

$$\sigma_{XY} = \frac{\sum (X_i - \bar{X})(Y_i - \bar{Y})}{n} = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{n} = \sum X_i Y_i - n \bar{X} \bar{Y} \quad (\text{B.4})$$

B.3.2: Coeficiente de correlação de Pearson (ρ):

Mede o grau de associação entre duas variáveis X e Y .

Definição:

$$\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} \quad (\text{B.5})$$

onde

σ_{XY} é a covariância de X e Y (dispersão conjunta)

σ_X é o desvio padrão de X (dispersão de X)

σ_Y é o desvio padrão de Y (dispersão de Y)

Propriedades:

- ρ é adimensional.
- ρ é limitado por 1, isto é, $|\rho_{XY}| \leq 1$.
- Valores de ρ_{XY} próximos de ± 1 indicam forte correlação entre as variáveis.

B.4: Testes de aderência:

Os testes de aderência são utilizados para verificar a adequabilidade de um modelo probabilístico para uma dada situação. Um desses testes usa a distribuição Qui-Quadrado, e é explanado a seguir.

B.4.1: Teste do qui-quadrado:

O teste qui-quadrado compara a frequência observada O_{ij} com a frequência esperada E_{ij} para cada célula. A estatística do teste é dada por:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \quad (\text{B.6})$$

O teste de aderência cria, então, o critério para decidir por aceitar ou não o modelo indicado. Em outras palavras, decidimos se os dados amostrais aderem ou não ao modelo. As hipóteses do teste são:

H_0 : X segue o modelo proposto.

H_a : X não segue o modelo proposto.

Nível de significância (α) = 0,05

Área de rejeição do teste:

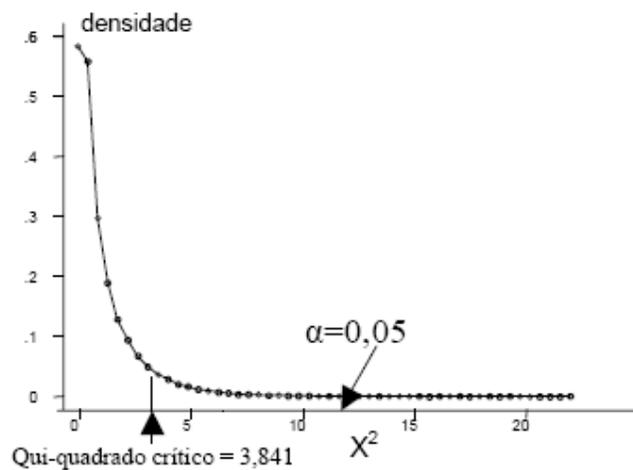
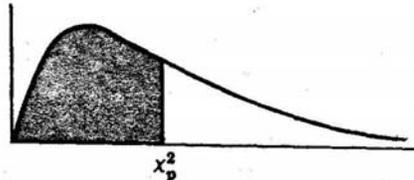


Figura B.1: Teste do qui-quadrado.

Decisão: rejeita-se H_0 se o valor calculado do qui-quadrado for maior do que o valor crítico, para um nível de significância pré definido.

VALORES DOS PERCENTIS (χ_p^2)
da
DISTRIBUIÇÃO DE QUI QUADRADO
com ν graus de liberdade
(área sombreada = p)



ν	$\chi_{0,995}^2$	$\chi_{0,99}^2$	$\chi_{0,975}^2$	$\chi_{0,95}^2$	$\chi_{0,90}^2$	$\chi_{0,75}^2$	$\chi_{0,50}^2$	$\chi_{0,25}^2$	$\chi_{0,10}^2$	$\chi_{0,05}^2$	$\chi_{0,025}^2$	$\chi_{0,01}^2$	$\chi_{0,005}^2$
1	7,88	6,63	5,02	3,84	2,71	1,32	0,455	0,102	0,0158	0,0039	0,0010	0,0002	0,0000
2	10,6	9,21	7,38	5,99	4,61	2,77	1,39	0,575	0,211	0,103	0,0506	0,0201	0,0100
3	12,8	11,3	9,35	7,81	6,25	4,11	2,37	1,21	0,584	0,352	0,216	0,115	0,072
4	14,9	13,3	11,1	9,49	7,78	5,39	3,36	1,92	1,06	0,711	0,484	0,297	0,207
5	16,7	15,1	12,8	11,1	9,24	6,63	4,35	2,67	1,61	1,15	0,831	0,554	0,412
6	18,5	16,8	14,4	12,6	10,6	7,84	5,35	3,45	2,20	1,64	1,24	0,872	0,678
7	20,3	18,5	16,0	14,1	12,0	9,04	6,35	4,25	2,83	2,17	1,69	1,24	0,989
8	22,0	20,1	17,5	15,5	13,4	10,2	7,34	5,07	3,49	2,73	2,18	1,65	1,34
9	23,6	21,7	19,0	16,9	14,7	11,4	8,34	5,90	4,17	3,33	2,70	2,09	1,73
10	25,2	23,2	20,5	18,3	16,0	12,5	9,34	6,74	4,87	3,94	3,25	2,56	2,16
11	26,8	24,7	21,9	19,7	17,3	13,7	10,3	7,58	5,58	4,57	3,82	3,05	2,60
12	28,3	26,2	23,3	21,0	18,5	14,8	11,3	8,44	6,30	5,23	4,40	3,57	3,07
13	29,8	27,7	24,7	22,4	19,8	16,0	12,3	9,30	7,04	5,89	5,01	4,11	3,57
14	31,3	29,1	26,1	23,7	21,1	17,1	13,3	10,2	7,79	6,57	5,63	4,66	4,07
15	32,8	30,6	27,5	25,0	22,3	18,2	14,3	11,0	8,55	7,26	6,26	5,23	4,60
16	34,3	32,0	28,8	26,3	23,5	19,4	15,3	11,9	9,31	7,96	6,91	5,81	5,14
17	35,7	33,4	30,2	27,6	24,8	20,5	16,3	12,8	10,1	8,67	7,56	6,41	5,70
18	37,2	34,8	31,5	28,9	26,0	21,6	17,3	13,7	10,9	9,39	8,23	7,01	6,26
19	38,6	36,2	32,9	30,1	27,2	22,7	18,3	14,6	11,7	10,1	8,91	7,63	6,84
20	40,0	37,6	34,2	31,4	28,4	23,8	19,3	15,5	12,4	10,9	9,59	8,26	7,43
21	41,4	38,9	35,5	32,7	29,6	24,9	20,3	16,3	13,2	11,6	10,3	8,90	8,03
22	42,8	40,3	36,8	33,9	30,8	26,0	21,3	17,2	14,0	12,3	11,0	9,54	8,64
23	44,2	41,6	38,1	35,2	32,0	27,1	22,3	18,1	14,8	13,1	11,7	10,2	9,26
24	45,6	43,0	39,4	36,4	33,2	28,2	23,3	19,0	15,7	13,8	12,4	10,9	9,89
25	46,9	44,3	40,6	37,7	34,4	29,3	24,3	19,9	16,5	14,6	13,1	11,5	10,5
26	48,3	45,6	41,9	38,9	35,6	30,4	25,3	20,8	17,3	15,4	13,8	12,2	11,2
27	49,6	47,0	43,2	40,1	36,7	31,5	26,3	21,7	18,1	16,2	14,6	12,9	11,8
28	51,0	48,3	44,5	41,3	37,9	32,6	27,3	22,7	18,9	16,9	15,3	13,6	12,5
29	52,3	49,6	45,7	42,6	39,1	33,7	28,3	23,6	19,8	17,7	16,0	14,3	13,1
30	53,7	50,9	47,0	43,8	40,3	34,8	29,3	24,5	20,6	18,5	16,8	15,0	13,8
40	66,8	63,7	59,3	55,8	51,8	45,6	39,3	33,7	29,1	26,5	24,4	22,2	20,7
50	79,5	76,2	71,4	67,5	63,2	56,3	49,3	42,9	37,7	34,8	32,4	29,7	28,0
60	92,0	88,4	83,3	79,1	74,4	67,0	59,3	52,3	46,5	43,2	40,5	37,5	35,5
70	104,2	100,4	95,0	90,5	85,5	77,6	69,3	61,7	55,3	51,7	48,8	45,4	43,3
80	116,3	112,3	106,6	101,9	96,6	88,1	79,3	71,1	64,3	60,4	57,2	53,5	51,2
90	128,3	124,1	118,1	113,1	107,6	98,6	89,3	80,6	73,3	69,1	65,6	61,8	59,2
100	140,2	135,8	129,6	124,3	118,5	109,1	99,3	90,1	82,4	77,9	74,2	70,1	67,3

Figura B.2: Tabela de valores do qui-quadrado.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)