
Sistemas especialistas probabilísticos e redes neurais na análise do crédito bancário

Maria Teresinha Arns Steiner
Celso Carnieri
Bruno H. Kopittke
Pedro J. Steiner Neto

As operações de crédito, como concessão de empréstimos, financiamentos, fianças, cartões de crédito e cheques especiais, compõem uma parte substancial das receitas de um banco comercial. Qualquer erro na decisão de concessão pode significar que em uma única operação haja a perda do ganho obtido em dezenas de outras bem-sucedidas. A correta decisão de crédito é essencial para a sobrevivência das empresas bancárias. É desejável e necessário, portanto, analisar uma proposta de negócio e comparar o **custo de conceder** com o **custo de negar** a operação.

A relação risco/retorno está implícita na concessão de crédito. O volume de incobráveis, assim como sua rentabilidade são efeitos da política adotada pela organização e de seus critérios de concessão de crédito. A otimização dos resultados é, portanto, decorrência de eficiente política de crédito, associada, evidentemente, à política de cobrança e às demais políticas da empresa (Silva, 1993).

A política de crédito de um banco comercial compreende o estabelecimento de taxas de juros, prazos, garantias e nível de risco relacionados a cada tipo de operação. A análise do processo decisório é bastante complexa, envolvendo, além da experiência anterior do analista do banco, instrumentos e técnicas que possam auxiliá-lo nessa tarefa.

Os métodos quantitativos, muito utilizados nesse tipo de análise, consideram registros históricos para a decisão sobre a concessão de crédito. Essas técnicas, entre as quais podem ser citadas Árvores de Decisão, Sistemas Especialistas e Redes Neurais, se empregadas corretamente, constituem eficientes ferramentas auxiliares dos gestores de crédito.

Decidida a concessão de crédito ao cliente, as três questões básicas que servem para orientar a definição do limite de crédito são (Rosenberg & Gleit, 1992; Curnow *et alii*, 1997):

- quanto o cliente merece de crédito (qualidade de risco apresentado e do **porte** do cliente);
- quanto pode ser oferecido de crédito ao cliente (capacidade de quem vai conceder o crédito);
- quanto se deve, de fato, conceder de crédito ao cliente (política de crédito adotada).

A autora agradece o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processo 300221/96-7.

Recebido em setembro/98
2ª versão em dezembro/98

Maria Teresinha Arns Steiner, Mestre e Doutora em Engenharia de Produção, na área de Pesquisa Operacional, pela Universidade Federal de Santa Catarina, é Professora Adjunta IV do Departamento de Matemática e Vice-Coordenadora do Curso de Mestrado em Métodos Numéricos em Engenharia da Universidade Federal do Paraná.

Fax: (41) 267-4236
E-mail: tere@mat.ufpr.br

Celso Carnieri, Mestre em Matemática pelo IMPA, Doutor em Engenharia Elétrica, na área de Otimização de Sistemas, pela Universidade Estadual de Campinas, Pós-Doutor pela Universidade de Illinois, Estados Unidos, é Professor Titular do Departamento de Matemática da Universidade Federal do Paraná.

Fax: (41) 267-4236
E-mail: carnieri@mat.ufpr.br

Bruno H. Kopittke, Mestre em Engenharia de Produção pela Universidade Federal de Santa Catarina, Doutor pelo Institut Polytechnique de Lorraine, França, Pós-Doutor pela Universität Der Bundeswehr Hamburg, Alemanha, é Professor Titular do Departamento de Engenharia de Produção da Universidade Federal de Santa Catarina e Professor Visitante da Universidade Federal do Paraná.

Fax: (41) 267-4236
E-mail: kopittke@eps.ufsc.br

Pedro J. Steiner Neto, Mestre em Engineering Management pelo Florida Institute of Technology, Flórida, Estados Unidos, Doutor em Administração pela Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, é Professor Adjunto IV do Departamento de Administração da Universidade Federal do Paraná e Diretor Faculdade de Administração e Economia (FAE), Curitiba, Paraná.

Fax: (41) 310-3195
E-mail: steiner@avalon.sul.com.br

As vantagens da utilização de técnicas quantitativas na administração de crédito são (Rosenberg & Gleit, 1992; Curnow *et alii*, 1997), entre outras:

- maior número de merecedores de crédito receberá o crédito (ou crédito adicional), aumentando os lucros;
- maior número de não merecedores de crédito terá o crédito negado (ou reduzido), diminuindo as perdas;
- os pedidos de crédito podem ser processados rapidamente;
- as decisões são objetivas e não passam por critérios subjetivos;
- menor número de pessoas é necessário para administrar o crédito e maior número com experiência pode concentrar-se nos casos mais difíceis.

O objetivo no presente trabalho é abordar duas técnicas da Pesquisa Operacional enquadradas na área da Inteligência Artificial, os Sistemas Especialistas Probabilísticos e as Redes Neurais, que se têm mostrado bastante promissoras na análise do crédito bancário visando discriminar (reconhecer, separar) os clientes adimplentes dos inadimplentes. Feito o reconhecimento desses padrões (fase da aprendizagem), essas técnicas podem ser utilizadas no diagnóstico de novos clientes (Fu, 1982).

DESCRIÇÃO DO PROBLEMA DO CRÉDITO BANCÁRIO

Os dados apresentados a seguir, e que serão utilizados no decorrer deste artigo, são reais e referem-se à concessão de empréstimos entre 20.000 e 50.000 DM por um banco alemão (Rödter, Kopittke & Kulmann, 1996).

Os critérios de avaliação do crédito utilizados neste trabalho são baseados em dados históricos, correspondentes a sete informações referentes a 2.855 clientes quanto ao seu desenrolar comportamental bancário, com respostas **sim** (1) ou **não** (0), e à informação fornecida pelo banco com relação ao cumprimento ou não do crédito, pelos mesmos, **pagou** (sim) ou **não pagou** (não), e comporão o banco de dados que será usado pelas técnicas aqui abordadas. Essas oito informações (de **A** a **G** e **K**), que constituirão as variáveis do problema analisado, são apresentadas a seguir.

A. Existência de rendimentos compatíveis? (sim/não) — (AREND)

Existem rendimentos compatíveis quando:

$$0,95Y - \max \{F, C\} - R \geq 0, \text{ onde:}$$

Y = rendimento mensal líquido;

C = gastos fixos (aluguéis, manutenção e outros);

F = limite de liberação de garantia;

R = prestações mensais referentes ao crédito.

B. Existência de patrimônio? (sim/não) — (BPATR)

É considerado o patrimônio compatível com a operação.

C. Estabilidade empregatícia há mais de três anos? (sim/não) — (CEMPR)

D. O solicitante é cliente? (sim/não) — (DCLIEN)

Este critério é preenchido quando o candidato apresenta uma conta corrente sem problemas há seis meses ou mais.

E. Inexistência de crédito insolvente pendente? (sim/não) — (EINSOL)

F. Inexistência de problemas no SPC? (sim/não) — (FSPC)

Qualquer tipo de problema conduz a uma resposta **não**.

G. Existência de fiador? (sim/não) — (GFIAD)

K. Crédito honrado? (sim/não) — (KPAGOU)

Foram considerados insolventes aqueles casos que assim podem ser lançados contabilmente. Não foram considerados atrasos de até um mês ou prorrogações de até três meses.

No quadro 1 constam as respostas para os 2.855 casos pesquisados. Das 128 (2⁷) possíveis combinações para as respostas às perguntas de A a G, só foram consideradas aquelas que efetivamente ocorreram.

Na terceira linha do quadro 1, por exemplo, lê-se o seguinte: dos dez clientes (KPAGOU_{total}) com AREND=0, BPATR=0, CEMP=1, DCLIEN=0, EINSOL=0, FSPC=0, GFIAD=0, oito casos tiveram um desenrolar positivo (KPAGOU=sim), enquanto em dois casos não foram honrados os pagamentos (KPAGOU=não).

Como pode ser observado, no quadro 1 as combinações de observações foram preenchidas em ordem crescente de ocorrência: primeiro os casos raros e depois os mais freqüentes.

Mesmo havendo a tendência de o número de observações positivas de clientes (respostas para as perguntas de A a G serem iguais a 1) correlacionar-se com um desenrolar positivo (KPAGOU=sim), esse fato não foi considerado neste trabalho.

Mesmo havendo a tendência de o número de observações positivas de clientes (respostas para as perguntas de A a G serem iguais a 1) correlacionar-se com um desenrolar positivo (KPAGOU=sim), esse fato não foi considerado neste trabalho.

SISTEMAS ESPECIALISTAS PROBABILÍSTICOS

Sistemas Especialistas são programas que têm a capacidade de adquirir e processar conhecimentos a partir das informações obtidas de um especialista em determinada

Quadro 1

Respostas da Amostra para os Casos de Concessão de Crédito Pesquisados

Número	A	B	C	D	E	F	G	KPAGOU		Total
	REND	PATR	EMP	CLIEN	INSOL	SPC	FIAD	Sim	Não	
1	0	0	0	0	0	0	0	8	2	10
2	0	0	0	0	0	0	1	8	2	10
3	0	0	1	0	0	0	0	8	2	10
4	1	0	1	0	0	0	0	8	4	12
5	1	0	0	0	0	0	1	17	2	19
6	0	0	0	0	0	1	0	17	2	19
7	0	0	1	0	0	0	1	17	3	20
8	1	0	0	0	0	0	0	17	3	20
9	1	1	0	0	0	0	0	17	5	22
10	1	0	0	1	1	1	0	26	2	28
11	1	0	0	0	0	1	1	26	2	28
12	1	0	1	0	0	1	0	26	2	28
13	1	0	1	1	0	1	0	26	2	28
14	0	0	0	1	0	1	0	26	3	29
15	1	0	0	0	0	1	0	26	3	29
16	1	1	1	0	0	0	0	26	4	30
17	1	1	1	1	0	1	0	34	1	35
18	1	0	0	1	0	1	0	34	2	36
19	1	1	1	0	0	1	0	34	2	36
20	1	1	1	1	1	1	0	34	2	36
21	0	0	0	1	1	1	0	34	3	37
22	0	0	1	0	0	1	0	34	4	38
23	1	1	0	0	0	0	1	34	5	39
24	1	0	0	1	0	1	1	43	1	44
25	1	0	0	1	1	1	1	43	2	45
26	0	0	0	0	0	1	1	43	3	46
27	0	0	1	1	0	1	0	43	3	46
28	1	1	1	0	0	0	1	43	4	47
29	1	0	1	0	0	0	1	43	7	50
30	1	1	1	1	0	1	1	52	1	53
31	1	1	0	1	1	1	0	52	2	54
32	0	0	1	1	1	1	0	52	3	55
33	1	0	1	1	1	1	0	60	2	62
34	0	0	1	0	0	1	1	60	3	63
35	0	0	0	1	0	1	1	64	2	66
36	0	0	0	1	1	1	1	78	2	80
37	1	1	0	0	0	1	0	78	2	80
38	1	1	0	1	0	1	0	86	3	89
39	0	0	1	1	1	1	1	86	4	90
40	1	0	1	0	0	1	1	95	2	97
41	1	1	0	0	0	1	1	95	3	98
42	0	0	1	1	0	1	1	95	4	99
43	1	1	0	1	1	1	1	121	2	123
44	1	1	0	1	0	1	1	130	2	132
45	1	1	1	0	0	1	1	156	3	159
46	1	0	1	1	1	1	1	173	2	175
47	1	1	1	1	1	1	1	190	2	192
48	1	0	1	1	0	1	1	208	3	211
Total								2.726	129	2.855

Fonte: Rödder, Kopittke & Kulmann (1996).

área. Quando consultados, esses programas fornecem apoio ao usuário tornando disponíveis os conhecimentos e as informações correspondentes (Rödder, Kopittke & Kulmann, 1996).

Já existem alguns programas de Sistemas Especialistas desenvolvidos e que vêm sendo muito utilizados, como, por exemplo, o MYCIN, desenvolvido na Universidade de Stanford, Estados Unidos, que faz o diagnóstico médico de infecções bacterianas e meningíticas.

Os estudos tiveram início na década de 70, mas somente a partir dos anos 90 houve destacado desenvolvimento, especialmente na área de sistemas especialistas probabilísticos. Um Sistema Especialista é denominado probabilístico (SEP) quando a base de conhecimento pode ser representada por uma distribuição de probabilidades, normalmente hiperdimensional (Kopittke, Dettmer & Gartner, 1996).

O SPIRIT, desenvolvido na FernUniversität, Alemanha, é uma interface de desenvolvimento de SEP para a construção e a aplicação de uma base de conhecimento sobre conjuntos de variáveis discretas (com seus atributos), cuja sigla incorpora suas principais características: *Symmetrical* (a inferência pode ser realizada nos dois sentidos, isto é, da premissa para a conclusão e vice-versa), *Probabilistic* (os prognósticos tomam como base a distribuição marginal de probabilidades), *Intentional* (a probabilidade dos fatos e regras que formam a estrutura de dependências entre as variáveis pode ser modificada por meio de observações no mundo real), *Reasoning* (envolve raciocínio lógico), *Inference Networks* (a inferência é efetivada a partir da transformação de um grafo em uma árvore de decisão) e *Transition* (as variáveis e suas relações podem estar em contínua mudança) (Kopittke, Dettmer & Gartner, 1996).

Conforme apresentado por Meyer & Rödder (1996), a base de conhecimento consiste em um conjunto finito de variáveis $V = \{V_1, \dots, V_n\}$ com domínio finito e uma distribuição de probabilidade P sobre o campo de todos os eventos em V . Os eventos são identificados com sentenças proposicionais $V_i = v_i$ (onde v_i é uma realização de V_i) construídas por negação, conjunção e disjunção. Se S é uma sentença desse tipo, então sua probabilidade é:

$$P(S) = \sum_{v \subset S} P(v) \quad [1]$$

onde v varia através de todas as conjunções completas de S na forma normal disjuntiva canônica de S .

Para gerar uma distribuição P que represente o conhecimento, primeiro estabelecem-se as probabilidades x_f e x_r aos fatos e regras, respectivamente. Um fato F é uma sentença proposicional, como definida anteriormente, e uma regra é uma expressão $F_2 | F_1$ com F_1 e F_2 sendo os fatos. As designações são entendidas como imperativas: ache uma P para a qual $P(F) = x_f$ ou $P(F_2 | F_1) = x_r$, cujas pro-

habilidades são não condicionadas e condicionadas, respectivamente.

Desse modo, para a aquisição de conhecimento no SPIRIT deve-se formular um conjunto de m fatos e regras, assim como estabelecer probabilidades consistentes:

$$P(F_{2i} | F_{1i}) = x_i, \quad i = 1, \dots, m \quad [2]$$

Em geral [2] não determina uma única P , mas um conjunto de distribuições. O conceito de entropia relativa auxilia a remover esse problema.

Para duas distribuições P e P_0 sobre V , a entropia relativa de P com respeito a P_0 é definida como:

$$R(P, P_0) = \sum_v P(v) \text{ld} [P(v) / P_0(v)] \quad [3]$$

com ld denotando o logaritmo na base dois.

Então, o problema consiste em resolver, utilizando os fatos e regras definidos, o seguinte problema para a aquisição do conhecimento:

$$\text{Min} \sum_v P(v) \text{ld} [P(v) / P_0(v)]$$

sujeito a: [4]

$$P(F_{2i} | F_{1i}) = x_i, \quad i = 1, \dots, m$$

No SPIRIT resolve-se o problema [4] anterior iterativamente, aplicando-se apenas uma restrição a cada iteração e, conseqüentemente, gerando-se uma seqüência de distribuições $P_0, P_1, \dots, P_{k-1}, P_k, \dots$ para as quais cada P_k tem uma entropia relativa mínima com respeito a P_{k-1} . Esse procedimento converge para uma distribuição ótima P^* .

Cada regra ou fato envolve certo grupo de variáveis, estabelecendo as relações de dependência entre elas. Então, cada conjunto de variáveis forma de maneira natural um *cluster* para o qual a distribuição de probabilidade deve ser satisfeita. O conjunto de todos os *clusters* constitui um hipergrafo em V , em geral cíclico. O SPIRIT, portanto, procura por um hipergrafo — uma hiperárvore —, de tal forma que cada um dos referidos *clusters*, como um todo, esteja contido em um hiperarco. Para esses hiperarcos as distribuições marginais são satisfeitas. Dessa forma, cálculos globais são reduzidos a cálculos locais: modificações em um hiperarco podem ser **propagadas** através de toda a hiperárvore.

O SPIRIT permite, ainda, que o aprendizado das regras seja do tipo alfa. O aprendizado alfa permite que a aprendizagem seja feita a partir de uma amostra de qualquer tamanho. Em cada hiperarco, toda célula é atualizada pela fórmula:

$$p_{\text{novo}} = (1 - \alpha)p_{\text{velho}} + \alpha f$$

onde p_{velho} é a probabilidade antes da aprendizagem alfa e f é a frequência na amostra. O valor de alfa, $0 \leq \alpha \leq 1$, é um peso que deve ser adequadamente escolhido pelo usuário. A base de conhecimento então aplica p_{novo} ao conjunto de regras que estejam previamente definidas e aplica o processo de aprendizagem novamente. A idéia do aprendizado alfa é ajustar probabilidades subjetivas por meio de probabilidades objetivas.

APLICAÇÃO DOS SISTEMAS ESPECIALISTAS PROBABILÍSTICOS AO PROBLEMA DO CRÉDITO BANCÁRIO

Conforme apresentado no tópico anterior, o processamento do conhecimento consiste na construção da hiperárvore, assim como no cálculo iterativo da distribuição conjunta para o problema.

Para construir a hiperárvore, o SPIRIT oferece algoritmos ótimos (Meyer & Rödder, 1996), cuja eficiência pode variar para diferentes bases de conhecimento. Assim que a hiperárvore fica determinada, o SPIRIT começa um processo iterativo aplicando todos os fatos e regras à regra empregada no momento. O SPIRIT detecta se a base de conhecimento está pronta ou se regras inconsistentes foram aplicadas pelo usuário.

A correta decisão de crédito é essencial para a sobrevivência das empresas bancárias.

Considerando as oito variáveis obtidas para cada um dos 2.855 clientes, apresentadas anteriormente, e definindo as sete primeiras como booleanas (com os atributos implícitos verdadeiro/falso, caso a resposta tenha sido sim ou não, respectivamente) e a última como nominal (com os atributos sim/não, caso o crédito tenha sido honrado ou não, respectivamente), foi proposta uma estrutura preliminar do espaço de observações. A decisão para o problema abordado neste trabalho foi a estrutura de interdependências apresentadas a seguir, definida juntamente com especialistas da área bancária:

- os dados sociais das observações **AREND**, **BPATR** e **CEMP**, de cada cliente, influenciam **KPAGOU**;
- as informações bancárias **DCLIEN**, **EINSOL** e **FSPC**, de cada cliente, atuam sobre **KPAGOU**;
- a existência de um fiador **GFIAD** sustenta **KPAGOU**, obtendo-se assim três *clusters*.

Para que o SPIRIT reconheça padrões de inter-relações em $\{\text{AREND, BPATR, CEMP, KPAGOU}\}$, $\{\text{DCLIEN, EINSOL, FSPC, KPAGOU}\}$ e $\{\text{GFIAD, KPAGOU}\}$ foram

Quadro 2

Fato e Regras para o Problema

Combinações	Tipo
AREND \wedge BPATR \wedge CEMP \Rightarrow KPAGOU=sim	Regra
AREND \wedge BPATR \wedge \neg CEMP \Rightarrow KPAGOU=sim	Regra
AREND \wedge \neg BPATR \wedge CEMP \Rightarrow KPAGOU=sim	Regra
AREND \wedge \neg BPATR \wedge \neg CEMP \Rightarrow KPAGOU=sim	Regra
\neg AREND \wedge BPATR \wedge CEMP \Rightarrow KPAGOU=sim	Regra
\neg AREND \wedge BPATR \wedge \neg CEMP \Rightarrow KPAGOU=sim	Regra
\neg AREND \wedge \neg BPATR \wedge CEMP \Rightarrow KPAGOU=sim	Regra
\neg AREND \wedge \neg BPATR \wedge \neg CEMP \Rightarrow KPAGOU=sim	Regra
DCLIEN \wedge EINSOL \wedge FSPC \Rightarrow KPAGOU=sim	Regra
DCLIEN \wedge EINSOL \wedge \neg FSPC \Rightarrow KPAGOU=sim	Regra
DCLIEN \wedge \neg EINSOL \wedge FSPC \Rightarrow KPAGOU=sim	Regra
DCLIEN \wedge \neg EINSOL \wedge \neg FSPC \Rightarrow KPAGOU=sim	Regra
\neg DCLIEN \wedge EINSOL \wedge FSPC \Rightarrow KPAGOU=sim	Regra
\neg DCLIEN \wedge EINSOL \wedge \neg FSPC \Rightarrow KPAGOU=sim	Regra
\neg DCLIEN \wedge \neg EINSOL \wedge FSPC \Rightarrow KPAGOU=sim	Regra
\neg DCLIEN \wedge \neg EINSOL \wedge \neg FSPC \Rightarrow KPAGOU=sim	Regra
GFIAD \Rightarrow KPAGOU=sim	Regra
\neg GFIAD \Rightarrow KPAGOU=sim	Regra
KPAGOU=sim	Fato

Nota: O símbolo \wedge tem o significado aditivo de e e o \neg tem o significado de negação não.

criadas 18 regras e um fato, já que é esse o número de combinações possíveis para a estrutura de interdependências definida, conforme pode ser observado no quadro 2. Não foram atribuídos valores de probabilidade, os quais serão definidos a seguir por meio do aprendizado alfa.

Assim, com as variáveis e as regras definidas anteriormente, fez-se o processamento do conhecimento no SPIRIT por meio da **Aprendizagem Alfa**, com $\alpha = 1$ (valor selecionado). O SPIRIT efetuou cerca de 10 mil iterações no modelo [4] apresentado, sendo que a partir da milésima iteração o valor da entropia relativa sofreu pouca mudança após cada alteração da distribuição de probabilidade, ou seja, foi atingido o ponto de equilíbrio, próximo de zero, para a entropia.

A representação da aquisição de informação foi feita, neste trabalho, por meio de um grafo de dependências não direcionado: para cada variável gera-se um vértice e, além disso, dois vértices estão ligados por um arco se as variáveis aparecem na mesma regra. Nesses vértices estão dispostos

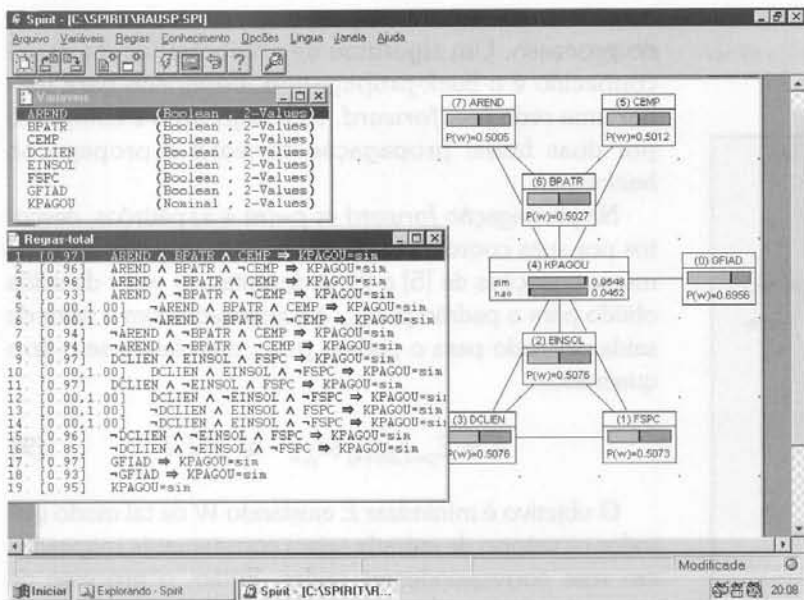


Figura 1: Representação do Conhecimento para o Problema Abordado — Situação Posterior ao Aprendizado

samento interconectadas. Em uma rede do tipo *feedforward* existem três tipos de unidades de processamento: de entrada, de saída e escondidas. As unidades de entrada recebem sinais do meio ambiente; as unidades de saída enviam sinais para o meio ambiente; as unidades escondidas não interagem diretamente com o ambiente, daí sua denominação, mas auxiliam no ajuste dos pesos da rede (Kröse & Van Der Smagt, 1993; Tam & Kiang, 1992).

O comportamento de cada unidade da rede pode ser modelado por funções matemáticas simples; conforme a rede neural ilustrada na figura 3, cuja topologia se enquadra no problema em questão de classificação dicotômica (apenas uma unidade de saída é necessária), uma unidade *i* recebe os sinais de entrada e os agrega com base em uma função de entrada:

rolamentos de barras que indicam a distribuição marginal atual da variável. A representação do conhecimento para o problema abordado aqui é apresentada na figura 1, correspondente à situação posterior ao aprendizado.

Feita a aprendizagem, o grafo de dependências constante na figura 1 pode ser usado para consultas. As perguntas ocorrem fixando-se o valor de uma ou mais variáveis aos valores desejados; obtém-se então, através da propagação dessa modificação temporária por toda a base de conhecimento, o valor da resposta (valor de **KPAGOU**) correspondente.

Na figura 2 mostra-se um exemplo do valor resposta — **KPAGOU** —, correspondente aos dez clientes da linha três do quadro 1, em que **AREND=0**, **BPATR=0**, **DCLIEN=0**, **EINSOL=0**, **FSPC=0**, **GFIAD=0**. O resultado do diagnóstico é, para esse caso, 71,35%.

De maneira análoga são obtidos os resultados do diagnóstico para todos os 2.855 clientes, cujas características se resumem aos 48 casos apresentados no quadro 1. Os resultados desse diagnóstico são mostrados no tópico **Análise dos Resultados**, na tabela 1.

REDES NEURAS

Uma rede neural (RN) consiste em determinado número de unidades de proces-

$$i_{p,i} = \sum_j w_{ij} x_{p,j} + \theta_i, \quad p = 1, \dots, (m+k) \quad [5]$$

$$i = 1, \dots, k^*$$

$$j = 1, \dots, n$$

onde:

- $i_{p,i}$ = entrada da unidade *i* para o padrão *p* (sendo que o número total de padrões é igual a *m+k*);
- i* = número de unidades na camada escondida;

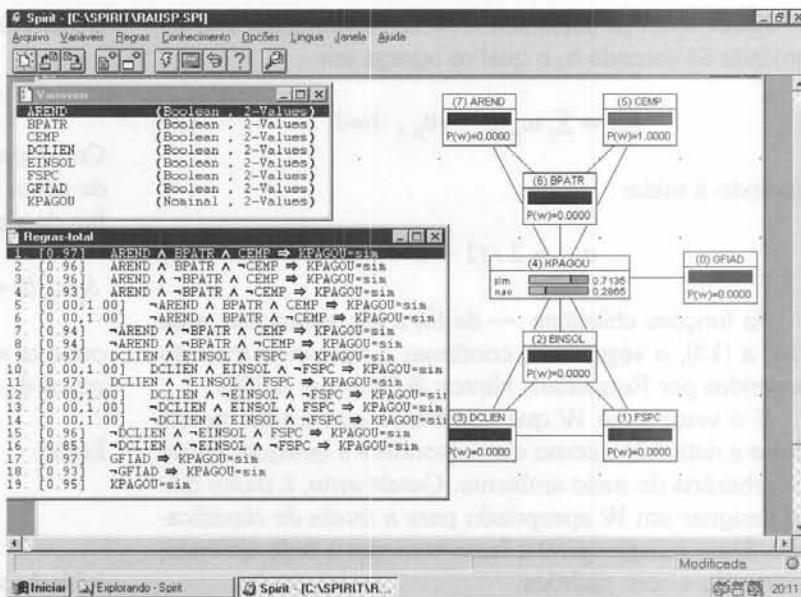


Figura 2: Exemplo do Valor Resposta — **KPAGOU**

w_{ij} = conexão peso entre as unidades i e j ;
 $x_{p,j}$ = entradas (coordenadas) do padrão p ;
 θ_i = bias da unidade i .

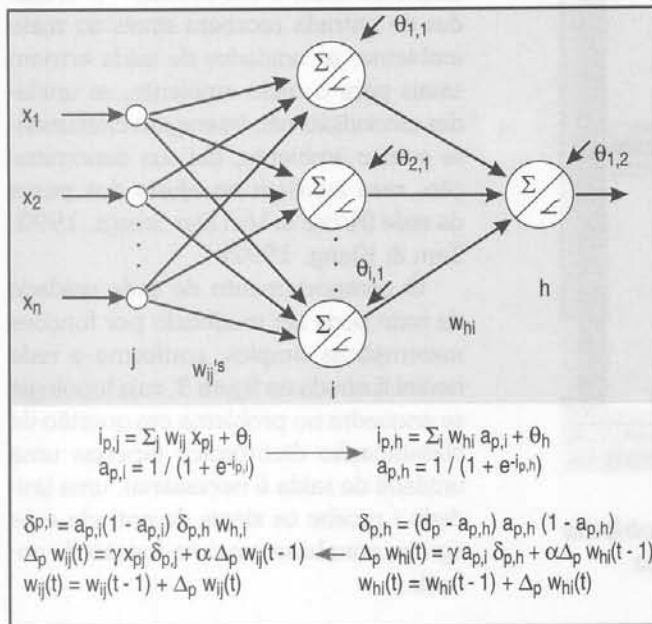


Figura 3: Ilustração do Algoritmo Back-Propagation em uma Rede Neural

Essa função de entrada gera um sinal de saída a_i para o padrão p , chamado de função de transferência:

$$a_{p,i} = 1 / (1 + e^{-i_{p,i}}) \quad [6]$$

Esses sinais de saída são então enviados para a única unidade da camada h , a qual os agrega em

$$i_{p,h} = \sum_i w_{hi} a_{p,i} + \theta_h, \quad h=1 \quad [7]$$

gerando a saída:

$$a_{p,h} = 1 / (1 + e^{-i_{p,h}}) \quad [8]$$

As funções utilizadas — de [5] a [8], anteriores, e de [10] a [13], a seguir —, contínuas e diferenciáveis, são sugeridas por Rumelhart, Hinton & Williams (1986).

É o vetor peso W que constitui o que a rede neural **sabe** e determina como ela responderá a qualquer entrada arbitrária do meio ambiente. Geralmente, é muito difícil designar um W apropriado para a tarefa de classificação. Uma solução geral é fazer com que a rede aprenda, treinando-a com padrões.

Um algoritmo de aprendizagem procurará através do espaço de W por um conjunto de pesos, oferecendo

o melhor ajuste para os pesos apresentados no início do processo. Um algoritmo de aprendizagem bastante conhecido é o *back-propagation*, designado para treinar uma rede *feedforward*. Esse algoritmo é composto por duas fases: propagação *forward* e propagação *backward*.

Na propagação *forward* os $p=(m+k)$ padrões, descritos por suas coordenadas $x_{p,j}$, alimentam a rede conforme as equações de [5] a [8] já descritas. O valor de saída obtido para o padrão p , $a_{p,h}$, é comparado com o valor de saída desejado para o padrão p , d_p , calculando-se o erro quadrático:

$$E = \sum_{p=1, (m+k)} (d_{p,h} - a_{p,h})^2 / 2 \quad [9]$$

O objetivo é minimizar E ajustando W de tal modo que todos os vetores de entrada sejam corretamente mapeados em suas correspondentes saídas. Então, o processo de aprendizagem pode ser visto como um problema de minimização com a função objetivo E definida no espaço de W .

A segunda fase, a propagação *backward*, que envolve as equações de [10] a [13] a seguir, executa um gradiente descendente em W para localizar a solução ótima. A direção e magnitude Δw_{ij} pode ser calculada da seguinte forma:

• **Variação em w_{hi} 's**

$$\Delta_p w_{hi} = -\gamma \partial E_p / \partial w_{hi}$$

onde γ = taxa de aprendizagem, $0 < \gamma < 1$.

Desenvolvendo as derivadas, chega-se em:

$$\Delta_p w_{hi} = \gamma a_{p,h} (1 - a_{p,h}) a_{p,i} (d_p - a_{p,h})$$

Considerando a situação atual t , para o padrão p , a troca de pesos ocorrida na situação $(t-1)$, para o padrão $(p-1)$, a fim de alcançar o mínimo mais rapidamente, tem-se:

$$\Delta_p w_{hi}(t) = \gamma a_{p,h} (1 - a_{p,h}) a_{p,i} (d_p - a_{p,h}) + \alpha \Delta_p w_{hi}(t-1) \quad [10]$$

onde α = constante que determina o efeito na troca de pesos da iteração $(t-1)$.

Então:

$$w_{hi}(t) = w_{hi}(t-1) + \Delta_p w_{hi}(t) \quad [11]$$

• **Variação em w_{ij} 's**

$$\Delta_p w_{ij} = -\gamma \partial E_p / \partial w_{ij}$$

Desenvolvendo as derivadas, chega-se em:

$$\Delta_p w_{ij} = \gamma (d_p - a_{p,h}) a_{p,h} (1 - a_{p,h}) w_{hi} a_{p,i} (1 - a_{p,i}) x_{p,j}$$

e, ainda,

$$\Delta_p w_{ij}(t) = \gamma (d_p - a_{p,h}) a_{p,h} (1 - a_{p,h}) w_{hi} a_{p,i} (1 - a_{p,i}) x_{p,j} + \alpha \Delta_p w_{ij}(t-1) \quad [12]$$

Então:

$$w_{ij}(t) = w_{ij}(t-1) + \Delta_p w_{ij}(t) \quad [13]$$

Uma ilustração geométrica desse método, com apenas duas variáveis e nenhum neurônio na camada escondida ($i = 0$), encontra-se na figura 4.

APLICAÇÃO DE REDES NEURAS AO PROBLEMA DO CRÉDITO BANCÁRIO

Para aplicar a técnica de Redes Neurais ao problema abordado, utilizou-se o pacote computacional *MatLab-Neural Network Toolbox* (Demuth & Beale, 1994), onde $m=2.726$ (KPAGOU=sim), $k=129$ (KPAGOU=não) e o número de características n é igual a oito (AREND, BPATR, CEMP, DCLIEN, EINSOL, FSPC, GFIAD e KPAGOU).

Para a topologia da Rede Neural optou-se por no máximo uma camada escondida com i unidades, onde $0 \leq i \leq k^*$, de acordo com a figura 3. A função de ativação utilizada na camada escondida e na de saída foi a função sigmoideal (Gorni, 1993).

Testes foram feitos para determinar o **melhor** número de neurônios da camada escondida, ou seja, que forneces-

se o menor erro. O número de neurônios i para essa camada escondida é determinado durante a execução do programa da seguinte forma: começa-se com $i=0$ e verifica-se o número de padrões classificados corretamente quando o processo de aprendizagem converge e, assim, prossegue-se até $i=k^*$. Dessas k^* tentativas escolhe-se, para i , aquela que classificou o maior número de padrões corretamente, denotada por i^* . Os pesos W e as bias θ , para cada uma dessas topologias $0 \leq i \leq k^*$, são arbitrários.

Todo o processo pode ser repetido para r conjuntos de pesos diferentes. Dessas r tentativas, registra-se aquela que, para a topologia com i^* unidades e com um conjunto de pesos W , apresentar melhor desempenho, ou seja, maior número de padrões classificados corretamente.

Vale salientar que a cada iteração na aplicação do algoritmo *back-propagation* é feito um sortimento nos $(m+k)$ padrões, ou seja, a cada iteração é modificado o seqüenciamento com que os $(m+k)$ padrões são apresentados para a rede neural.

Como regra de parada para cada uma das situações, em que o número de neurônios na camada escondida é $0 \leq i \leq k^*$, e conjunto de pesos W , adotou-se a seguinte alternativa: para cada uma das topologias pesquisadas em que a Rede Neural apresentar uma variação de erro entre duas iterações consecutivas menor do que um valor ϵ , finaliza-se o processo de aprendizagem e, então, toma-se o W , da topologia com i neurônios, que tiver apresentado maior porcentagem de acerto, i^* . Repete-se esse processo para os r conjuntos de pesos, obtendo-se r valores para i^* . Entre os r conjuntos de pesos, escolhe-se o conjunto de pesos W , correspondente a i^* , que classifique o maior número de padrões corretamente, finalizando assim todo o processo.

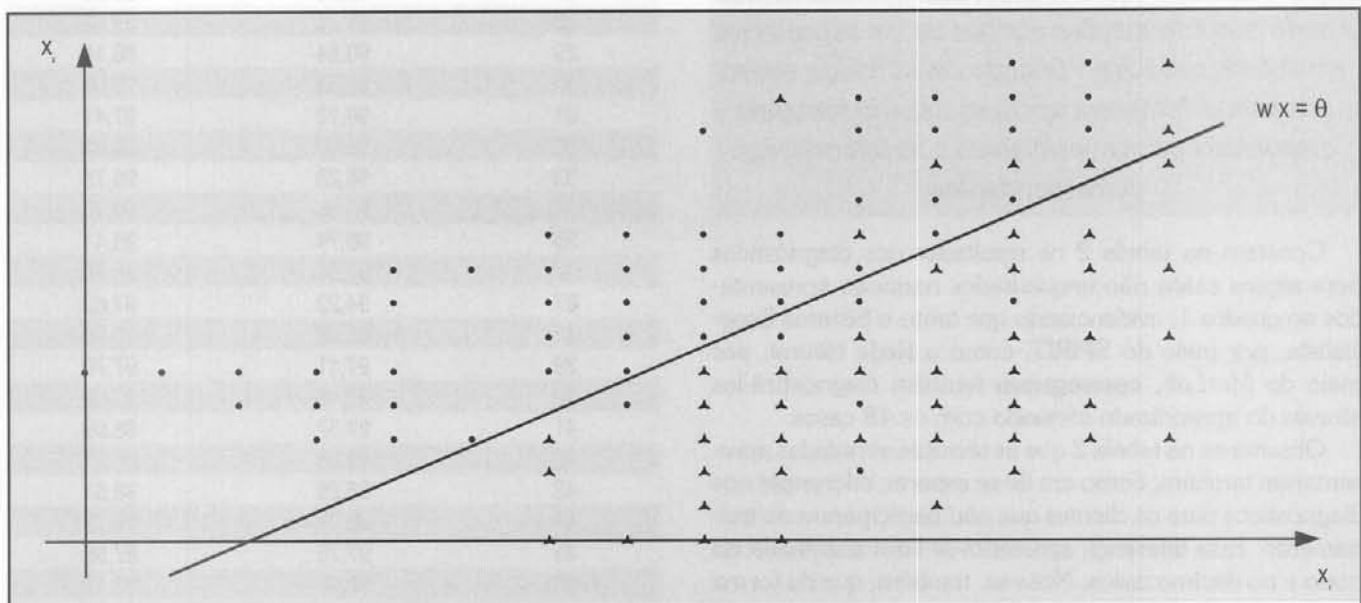


Figura 4: Ilustração Geométrica de uma Rede Neural

Pelo procedimento acima descrito (Steiner, 1995), para o problema abordado foi escolhida uma topologia composta por oito neurônios na camada escondida, a qual apresentou a menor porcentagem de erros na classificação dos padrões. Para essa técnica também foram efetuadas 10 mil iterações (uma iteração corresponde a uma passagem dos $(m + k)$ padrões através da rede) utilizando o *MatLab*, sendo que o valor do erro sofreu alterações consideráveis no início da aprendizagem, mas a partir da quingentésima iteração, aproximadamente, ele pouco se modificou.

Encerrada a aprendizagem, pôde-se obter os resultados dos diagnósticos para todos os 2.855 clientes enquadrados nos 48 casos mostrados no quadro 1. Esses resultados constam na tabela 1, permitindo compará-los com os alcançados por meio do Sistema Especialista Probabilístico.

ANÁLISE DOS RESULTADOS

Na tabela 1, pode-se observar e comparar os diagnósticos fornecidos pelas técnicas pesquisadas, Sistemas Especialistas Probabilísticos e Redes Neurais. Nessa tabela tem-se que, por exemplo, para os dez clientes da linha três do quadro 1 o Sistema Especialista diagnostica que o cliente pagará o seu crédito com uma probabilidade de 71,35%, fornecendo a Rede Neural uma probabilidade de 78,93% para esse caso. Com exceção das probabilidades dos casos contidos nas linhas quatro e oito, as demais são bastante próximas para as duas técnicas.

Sistemas Especialistas são programas que têm a capacidade de adquirir e processar conhecimentos a partir das informações obtidas de um especialista em determinada área. Quando consultados, esses programas fornecem apoio ao usuário tornando disponíveis os conhecimentos e as informações correspondentes.

Constam na tabela 2 os resultados dos diagnósticos para alguns casos não enquadrados naqueles apresentados no quadro 1, evidenciando que tanto o Sistema Especialista, por meio do SPIRIT, como a Rede Neural, por meio do *MatLab*, **conseguem** também diagnosticá-los através do aprendizado efetuado com os 48 casos.

Observa-se na tabela 2 que as técnicas abordadas apresentaram também, como era de se esperar, diferenças nos diagnósticos para os clientes que não participaram do treinamento. Essa diferença apresenta-se bem acentuada no nono e no décimo casos. Nota-se, também, que de forma geral os diagnósticos são mais baixos para a Rede Neural do que para o Sistema Especialista Probabilístico.

Tabela 1
Diagnóstico de Crédito — SPIRIT (SEP) e MatLab (RN)

Número	Diagnóstico SEP (%)	Diagnóstico RN (%)
1	70,57	82,18
2	84,50	80,93
3	71,35	78,93
4	81,36	69,79
5	83,57	88,70
6	90,32	86,14
7	84,99	84,90
8	69,11	83,48
9	80,72	78,65
10	92,98	94,54
11	95,19	95,76
12	94,44	91,51
13	95,96	94,11
14	92,89	90,50
15	89,69	89,01
16	83,08	85,83
17	96,39	96,79
18	92,41	93,35
19	95,02	95,00
20	96,68	93,98
21	93,42	90,74
22	90,64	91,14
23	90,49	87,53
24	96,51	96,92
25	96,79	96,90
26	95,50	93,00
27	93,13	93,96
28	91,78	89,62
29	90,84	86,16
30	98,38	99,06
31	96,12	97,41
32	96,80	94,81
33	96,28	96,75
34	95,66	96,46
35	96,74	95,41
36	97,00	96,78
37	94,22	97,62
38	95,96	96,50
39	97,11	97,70
40	97,48	97,06
41	97,37	98,08
42	96,86	96,51
43	98,26	98,61
44	98,11	98,72
45	97,75	97,98
46	98,33	98,19
47	98,51	98,66
48	98,18	98,56

Tabela 2

Diagnóstico de Combinações de Observações Não Aprendidas

Número	A REND	B PATR	C EMP	D CLIEN	E INSOL	F SPC	G FIAD	Diagnóstico SEP (%)	Diagnóstico RN (%)
1	0	1	1	0	0	0	0	83,64	76,72
2	0	1	1	1	1	1	1	98,57	98,54
3	0	0	0	0	1	1	1	96,80	96,32
4	1	1	1	0	1	1	1	98,41	88,91
5	1	1	1	1	1	0	1	98,41	85,62
6	0	0	0	0	0	1	0	90,32	86,14
7	0	0	0	1	0	0	0	93,01	86,92
8	0	1	0	1	0	1	0	96,53	93,92
9	1	0	1	0	1	0	1	98,22	71,81
10	0	0	1	1	1	0	0	93,25	63,17

Obtidos os diagnósticos, deve-se, agora, tomar a decisão sobre a concessão ou não do crédito a cada um dos diversos casos de clientes. Os administradores bancários (Souza & Clemente, 1997) sabem que essa decisão ainda depende, individualmente, de outros fatores, entre os quais:

- valor do crédito solicitado (VCS);
- taxa de juros de empréstimos;
- prazo;
- taxa bancária de abertura de crédito;
- custo do capital.

Para melhor entendimento, far-se-á o estudo considerando o exemplo: valor de crédito de R\$ 1.000,00, a uma taxa de empréstimos (taxa cobrada do cliente) de 5% ao mês, um custo de capital (taxa paga pelo banco) de 2% ao mês e considerando, ainda, uma taxa de abertura de crédito (TAC) de 3%, durante um prazo de 12 meses.

- O retorno bancário para um desenrolar positivo será de:
 - é dado ao cliente no início da operação o valor de 1.000 - 3% (1.000,00) = 970,00, resultando em ganho bancário inicial de R\$ 30,00;

- o valor do pagamento mensal PMT (payment) é de

$$PMT = (VCS) \frac{[i_{cliente} (1+i_{cliente})^n]}{[(1+i_{cliente})^n - 1]} \quad [14]$$

$$PMT = 1.000 \frac{[0,05(1+0,05)^{12}]}{[(1+0,05)^{12} - 1]}$$

$$PMT = 112,83$$

- o valor presente VP do empréstimo é de

$$VP = 30 + \frac{[(1+i_{banco})^n - 1]}{[i_{banco} (1+i_{banco})^{12}]} PMT - 970 \quad [15]$$

$$VP = 30 + \frac{[(1+0,02)^{12} - 1]}{0,02(1+0,02)^{12}} 112,83 - 1000$$

$$VP = 223,17$$

- então, o retorno total esperado é de R\$ 223,17, ou seja, de 22,32%, aproximadamente.

- O retorno (negativo) bancário para um desenrolar negativo será de:
 - é dado ao cliente no início da operação o valor de 1.000 - 3% (1.000,00) = 970,00, resultando em ganho bancário inicial de R\$ 30,00;
 - o retorno (negativo), considerando o valor inicial de R\$ 1.000,00 e prevendo o ressarcimento de 30% ao final do período, será de

$$Retorno = \text{Valor Ressarcido} / (1 + i_{banco})^n$$

$$Retorno = 300,00 / (1+0,02)^{12}$$

$$Retorno = 236,55$$

- então, o retorno (negativo) total será de R\$ 30,00 + R\$ 236,55 - R\$ 1.000,00 = -R\$ 733,45, ou seja, de 73,35%, aproximadamente.

Assim, a concessão de um crédito é interessante quando o resultado de:

$$p (VCS) 0,2232 - (1 - p) (VCS) 0,7335 \geq 0 \quad [16]$$

onde p = probabilidade de um desenrolar positivo.

Nas condições apresentadas (taxas e prazos), obtém-se por meio da equação [16] $p^* = 0,7668$, isto é, o crédito deve ser concedido para os clientes cuja probabilidade de adimplência seja igual ou superior a 76,68%; caso contrário, o crédito deve ser negado.

Considerando o valor obtido para p^* e comparando-o com os dados contidos na tabela 1, observa-se que em relação aos diagnósticos fornecidos:

- pelo Sistema Especialista Probabilístico, o crédito seria negado para os clientes enquadrados nos casos um, três e oito, e concedido para todos os demais clientes;
- pela Rede Neural, o crédito seria negado apenas para os clientes do caso quatro.

Analisando as equações [14] e [15] verifica-se que quanto maior o prazo (n) para pagamento, considerando-se as mesmas taxas, menor é o valor para p^* , ou seja, menores são as exigências com relação ao diagnóstico.

CONCLUSÕES

Neste trabalho, tem-se como objetivo apresentar duas

metodologias para o reconhecimento de padrões de comportamento de clientes com relação à adimplência de crédito, examinando o caso concreto de um banco alemão, e com base nesse reconhecimento prever o comportamento de futuros clientes, buscando minimizar as perdas bancárias e o esforço gerencial na decisão de concessão de crédito.

Tendo por base dados históricos, foram determinadas as probabilidades de adimplência dos 2.855 clientes pesquisados, utilizando-se as técnicas de Redes Neurais e Sistemas Especialistas Probabilísticos, comparativamente. Feito o **treinamento** dos programas computacionais envolvendo essas duas técnicas, pôde-se obter o diagnóstico para outros clientes não incluídos na amostra.

Obtidos os diagnósticos, foi feito um estudo mostrando em qual situação é interessante para o banco fazer a concessão de crédito, considerando o valor do crédito solicitado, a taxa de juros de empréstimos, o prazo, a taxa bancária de abertura de crédito e o custo do capital, de forma a maximizar os lucros e minimizar os riscos bancários. ♦

RESUMO

Reconhecer e prever quais clientes serão "bons ou maus pagadores" de crédito é tarefa importante e difícil para as instituições bancárias e os serviços de proteção ao crédito. Com registros históricos de 2.855 clientes de um banco alemão, foram abordadas neste artigo, comparativamente, as técnicas de Sistemas Especialistas Probabilísticos e Redes Neurais, ambas da área de Inteligência Artificial, utilizando o *SPIRIT* e o *MatLab-Neural Networks Toolbox*. Essas técnicas permitem o reconhecimento de padrões e, também, a sua utilização em diagnósticos posteriores. Feito o reconhecimento de determinado cliente e obtida a sua conseqüente probabilidade de retorno do crédito, estudou-se, ainda, com base nas taxas de juros adotadas, no valor do crédito a ser concedido, no prazo para pagamento e nas demais taxas bancárias, a relação risco/retorno para o banco, apontando os casos em que o crédito deve ser concedido.

Palavras-chave: reconhecimento de padrões, sistemas especialistas probabilísticos, redes neurais.

ABSTRACT

Recognising and foreseeing which credit customers will be "good or bad payers" is an important and difficult task to banking institutions and to credit protection services. With 2.855 historical registers of a German bank, we studied, in this work, the Probabilistic Expert Systems and Neural Networks techniques, both in the Artificial Intelligence area, comparatively, using the shell *SPIRIT* and *MatLab-Neural Networks Toolbox*. These techniques allow us to recognise patterns as well as their further use in later evaluations. The evaluation of a given customer and the probability of the payment of a loan leads to a risk/payment rate for the bank, based upon interest rates, loan amount, payment period and other fees, establishing a safe criteria to determine when to concede credit.

Uniterms: pattern recognition, probabilistic expert systems, neural networks.

- CURNOW, G.; KOCHMAN, G.; MEESTER, S.; SARKAR, D.; WILTON, K. Automating credit and collections decisions at AT&T capital corporation. *Interfaces*, v.27, p.29-52, 1997.
- DEMUTH, H. & BEALE, M. *Neural network toolbox for use with MATLAB (user's guide)*. Natick, Massachusetts, The Math Works, Inc., 1994.
- FU, K.S. *Syntatic pattern recognition and applications*. New Jersey, Prentice-Hall, 1982.
- GORNI, A.A. *Redes neurais artificiais — uma abordagem revolucionária em inteligência artificial*. São Paulo, Micro Sistemas, 1993.
- KOPITKE, B.H.; DETTMER, A.L.; GARTNER, I.R. Um sistema de apoio à decisão estratégica baseado em inferência probabilística. CONGRESSO NACIONAL DE INVESTIGAÇÃO OPERACIONAL, 7. *Anais*. Aveiro, Portugal, 1996.
- KRÖSE, B.J.A. & VAN DER SMAGT, P.P. *An introduction to neural networks*. Amsterdam, University of Amsterdam, 1993.
- MEYER, C.-H. & RÖDDER, W. Probabilistic knowledge representation and reasoning at maximum entropy by SPIRIT. ANNUAL GERMAN CONFERENCE ON ARTIFICIAL INTELLIGENCE, 20. *Proceedings*. Dresden, Germany, 1996.
- RÖDDER, W.; KOPITKE, B.; KULMANN, F. *Sistemas especialistas probabilísticos*. Texto para o Ensino à Distância feito em Cooperação com a Fern Universität Hagen, 1996.
- ROSENBERG, E. & GLEIT, A. Quantitative methods in credit management: a survey. *Operations Research*, v.42, n.4, p.589-613, 1994.
- RUMELHART, D.E.; HINTON, G.; WILLIAMS, R. Learning representation by back-propagation errors. *Nature*, v.323, n.9, p.533-536, 1986.
- SILVA, J. PEREIRA. *Análise e decisão de crédito*. São Paulo, Atlas, 1993.
- SOUZA, A. & CLEMENTE, A. *Decisões financeiras e análise de investimentos*. São Paulo, Atlas, 1997.
- STEINER, M.T.A. *Uma metodologia para o reconhecimento de padrões multivariados com resposta dicotômica*. Santa Catarina, 1995. Tese (Doutorado) em Engenharia de Produção — Universidade Federal de Santa Catarina.
- TAM, K.Y. & KIANG, M.Y. Managerial applications of neural networks: the case of bank failure predictions. *Management Sciences*, v.38, n.7, p.926-947, 1992.

II Gestão Empresarial de Qualidade de Vida

Curso Avançado

Setembro a Dezembro de 1999

PROGEP – Programa de Estudos em Gestão de Pessoas

FIA-FEA/USP

Será realizado na Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo (FEA/USP), pelo PROGEP, o **II Curso Avançado de Gestão Empresarial de Qualidade de Vida**, com 60 horas de duração (terças-feiras das 18h30 às 22h30), de 14 de setembro a 14 de dezembro de 1999, cujos objetivos são:

- Propiciar o desenvolvimento de pensamento crítico favorável à implementação de visão estratégica de Qualidade de Vida no Trabalho.
- Gerar condições de especialização profissional por meio de conceitos, modelos de gestão e ações práticas de Gestão Empresarial em Qualidade de Vida.
- Criar e fortalecer valores culturais com ênfase multidisciplinar para Gestão Empresarial da Qualidade de Vida.

Este curso é dirigido a empresários, executivos, pesquisadores e consultores com interesse em Qualidade de Vida e áreas afins, como gestão de pessoas, medicina, segurança do trabalho, promoção da saúde, serviço social e estratégias empresariais.

Maiores informações:



pelos telefones (11) 818-5835 – Raquel/PROGEP, (11) 3742-5982 – Zuleide, (19) 253-5866 – Mariângela



por correio eletrônico nos seguintes endereços: progep@fia.fea.usp.br ou cageqv@fia.fea.usp.br