

Avaliação de Desempenho Corporativa: o emprego de inteligência artificial na indústria bancária

Gustavo José de Guimarães e Souza

Mestre em Economia pela Universidade Federal Fluminense e Assessor do Banco do Brasil

RESUMO

A importância do processo avaliatório em nossa sociedade torna-se cada vez mais importante sob vários aspectos. No entanto, a análise de desempenho é passível de discussões. Neste sentido o artigo propõe colaborar. Pelo emprego de técnicas de inteligência computacional, no caso específico Lógica Nebulosa e Redes Neurais, sugerem-se dois tipos de modelos de avaliação de performance bancária: um modelo fuzzy e um modelo neuro-fuzzy. Modelos estes capazes de gerar uma nota de avaliação ponderada pelo trinômio rentabilidade-eficiência-risco. Os dados pesquisados referem-se ao sistema bancário brasileiro no ano de 2005, assim a análise multidimensional utilizada é relativa por considerar o desempenho individual de um banco dado o desempenho dos demais bancos brasileiros. A amostra compreende 97 instituições bancárias, cerca de 84,34% de todo o Sistema Financeiro Nacional no ano em questão.

Palavras-Chave: Bancos, Desempenho Organizacional, Lógica Fuzzy, Neuro-Fuzzy, Inteligência Artificial.

1. Introdução

Após o advento do Plano Real (julho de 1994), o sistema bancário brasileiro teve que se adaptar a uma nova realidade macroeconômica. O tradicional efeito inflacionário que ocultava as deficiências na condução de uma instituição bancária, já não mais existia. As melhorias de desempenho interno e externo tornam-se cruciais para manutenção das margens de lucro e da própria sobrevivência do banco no mercado. Logo, os processos de medição, de avaliação e de acompanhamento da performance bancária, tomam relevância sem precedências.

Apesar da importância de se avaliar o desempenho interno das unidades que compõem um banco, este não é o escopo do presente trabalho, mas sim o desempenho corporativo. Atualmente, o acompanhamento da performance bancária, e a conseqüente transparência no setor, permeiam vários aspectos econômicos. Bancos são instituições singulares, seu desempenho é importante não só para acionistas, funcionários, concorrentes, reguladores e clientes, mas para a economia como um todo. A capacidade de criação de moeda e a integração com o sistema de pagamentos permitem que uma instituição em dificuldades provoque um choque que possa ser transmitido ao sistema econômico como um todo, o chamado risco sistêmico. Desta forma, os demais entes da sociedade são como *stakeholders*¹ dos bancos.

O intento do presente estudo consiste em propor modelagem via inteligência computacional para avaliar o desempenho das instituições bancárias no Brasil. Os modelos permitem avaliação relativa das instituições no Brasil e paralelamente servem como ferramentas de auxílio para analistas de avaliação corporativa. Além desta introdução o artigo é composto de mais quatro seções. A segunda seção discorre brevemente sobre o

¹ O termo *stakeholders* é designado para pessoas ou empresas que, de alguma maneira, são influenciadas pelas ações de uma organização.

estado da arte na análise de desempenho corporativo apresentando os indicadores utilizados no trabalho e a base de dados consultada. A terceira seção propõe uma análise do desempenho bancário por via da lógica *fuzzy*. A quarta seção apresenta um modelo de avaliação de desempenho bancário por um sistema híbrido, o *neuro-fuzzy*. Por último é apresentada a conclusão.

2. Análise de Desempenho

2.1. INDICADORES CORPORATIVOS

Avaliar significa diferenciar. Assim, modelos de avaliação só surgem e se impõem quando os interesses a favor da diferenciação superam os interesses dos que preferem a igualdade indiferenciada (Schwartzman, 1986). O porquê se avaliar bancos se justifica pela própria relevância deste segmento na esfera econômica e social. Entretanto, ainda duas questões se colocam: o que medir e como mensurar?

Os indicadores são instrumentos que diagnosticam a situação da empresa e possibilitam a tomada de decisões. No entanto, diversos são os utilizados para verificar e analisar o desempenho bancário. A escolha de indicadores adequados parte do pressuposto que se devem definir medidas que de fato avaliem o que se pretende medir. Até a década de 80, predominavam no meio empresarial os indicadores de caráter puramente financeiro, tais como lucro, retorno sobre vendas, retorno sobre investimentos, venda por empregado, variação de preços e principalmente a produtividade (Miranda e Silva, 2002). Hoje, a afirmação de que o lucro, por exemplo, é a melhor e mais consistente medida de performance de uma organização e garante a continuidade desta, não convence.

No Brasil existem alguns anuários especializados que oferecem *rankings* de bancos. Entre eles: Maiores Bancos Comerciais e Múltiplos do Brasil (Fundação Getúlio Vargas); Melhores & Maiores (Revista Exame/Fipecaf-USP) e Valor 1000 (Jornal Valor Econômico). Contudo, estes *rankings* consideram indicadores em valor absoluto, ocasionando uma listagem decrescente de maiores bancos, mas não de bancos com melhor desempenho. Este tem caráter relativo. O desempenho corporativo de um banco aqui sugerido é consubstanciado em três pilares básicos: rentabilidade, produtividade e risco.

Apesar de relevante, a rentabilidade em determinado período não reflete toda a realidade da instituição se não conjugada com o risco que a mesma assumiu para atingi-la. A metodologia RAROC – *Risk Adjusted Return on Capital*, proposta pelo *Bankers Trust*, no final da década de 70, introduziu a concepção evolutiva nas métricas de avaliação do desempenho em bancos, não só orientada ao retorno observado sobre um referencial contábil ou apreçado a mercado, mas principalmente ao risco. Risco este incorrido nos negócios que originaram a rentabilidade observada². O terceiro aspecto a ser considerado é a produtividade, os processos internos, ou seja, como o banco transforma seus insumos em produtos e serviços financeiros.

Para cada dimensão escolhe-se o indicador mais representativo³, fundamentado na literatura específica e na consulta a especialistas de avaliação de desempenho, sendo todos os indicadores de caráter relativo. O RAROC que poderia responder por duas dimensões (rentabilidade e risco) não pôde ser adotado, pois para seu cálculo é necessário o valor do V@R (valor em risco). No entanto, existem restrições ao uso desta variável para os fins desejados aqui. Primeiro, nem todos os bancos a calculam; segundo, cada banco pode apurar

² Para maiores informações sobre RAROC ler Falkenstein (1997) e James (1996).

³ A escolha de um indicador representativo se dá pela simplificação metodológica, nada obstante, o uso de vários indicadores em cada dimensão, ou mesmo a utilização da análise fatorial para conjugar os possíveis indicadores existentes na literatura em *scores* para cada dimensão pode ser efetuada em trabalhos futuros.

seu cálculo por metodologias próprias não havendo uniformidade de informações; e por último, não é disponível publicamente para os casos onde é mensurada. No entanto, como será visto à frente, a consideração da rentabilidade dado o risco é mantida na análise pelo uso dos indicadores de rentabilidade e de risco tomados conjuntamente no modelo proposto.

Cada dimensão foi então contemplada por um único indicador. Para medição da rentabilidade o indicador sugerido é o Retorno sobre o Patrimônio Líquido Médio (RSPL). Este indica quanto o banco obteve de lucro para cada 100 unidades monetárias de capital próprio investido. Tem-se então a taxa de rendimento do capital. Sua interpretação é quanto maior melhor.

$$RSPL(\%) = \frac{\text{Lucro Líquido Acumulado}}{\text{Patrimônio Líquido Médio}} \quad (1)$$

Dentre os indicadores de produtividade o escolhido para o estudo é o índice de eficiência, por ser mais abrangente que os demais existentes na literatura (cobertura, despesas por funcionários, entre outros). Ele fornece o quanto das receitas obtidas está sendo consumido pelas despesas. Sua interpretação é quanto menor melhor.

$$\text{Eficiência (\%)} = \frac{DA + ODA}{RBIF - PCLD + RPS + RPCC + ORO + ODO} \quad (2)$$

onde: DA = despesas administrativas; ODA = outras despesas administrativas; RBIF = receita bruta de intermediação; PCLD = provisão para créditos de liquidação duvidosa; RPS = receita sobre prestação de serviços; RPCC = resultado de participações em coligadas e controladas; ORO = outras receitas operacionais; e ODO = outras despesas operacionais.

Atualmente não temos uma medida unificada de risco para as atividades bancárias. Os riscos de crédito, mercado e operacional⁴ não podem ainda serem integrados em uma medida singular (apesar dos esforços teóricos para realização de tal cálculo). Todavia, a relevância do risco de crédito no tocante a alocação de capital e a característica tradicional de intermediação financeira dos bancos, o possibilita ser uma *proxy* para o risco geral da instituição. O risco de crédito é representado aqui pela razão entre o valor provisionado em risco e a carteira total de crédito de um banco⁵. Sua interpretação é como do índice de eficiência: quanto menor o risco, melhor.

$$\text{Risco de Crédito (\%)} = \frac{PCLD}{\text{Carteira de Crédito}} \quad (3)$$

Os três indicadores de desempenho conjugados avaliam o banco sob três perspectivas diferentes, rentabilidade-eficiência-risco. São representativos também das esferas propostas por Matarazzo (1997): financeira e econômica.

2.2. DADOS

Todos os dados necessários ao cálculo dos indicadores são extraídos do Banco Central do Brasil -BCB (2006), visando à uniformidade e qualidade de procedência. As informações são divulgadas trimestralmente e obtidas junto aos próprios bancos por meio dos balancetes e balanços. O período de análise corresponde ao ano de 2005.

Para o cálculo do RSPL para cada banco são obtidos o lucro líquido acumulado no ano de referência e o patrimônio líquido médio que representa a soma do patrimônio do ano de 2004 com o de 2005, dividido por dois. Esta é uma forma de absorver as mudanças ocorridas

⁴ Apesar dos bancos estarem sujeitos a vários outros tipos de risco, esses são os principais.

⁵ Esta medida é sugerida por ser capaz de oferecer uma expectativa do quanto de sua carteira de crédito o banco espera que se torne inadimplente e/ou este pretende se precaver desta situação.

no capital investido durante o período de avaliação. Por sua vez, as informações referentes às receitas e despesas que compõem o índice eficiência são acumuladas para 2005. Quanto à análise do risco, o ideal seria a média diária do risco envolvido nas operações de crédito. Contudo, estas informações não são de acesso público, apesar de desejável sob a ótica da transparência e *disclosure*⁶. Então, risco medido aqui se refere à média dos quatro trimestres do ano em questão. O valor provisionado é constituído com base na resolução 2.682 do Conselho Monetário Nacional (CMN).

Dos 104 bancos pertencentes ao consolidado bancário I estabelecido pelo BCB (instituições independentes e conglomerados financeiros que aceitam depósitos à vista), 97 constituem a amostra. Dos sete bancos não presentes na amostra três são excluídos por apresentarem denominador negativo para o índice de eficiência (as receitas não cobrem nem mesmo o resultado negativo da participação em coligadas e/ou as ODA). Os quatro bancos restantes não possuem carteira de crédito, e portanto, não são passíveis de mensuração do risco via *proxy* adotada. A amostra, no entanto, tem representatividade considerável do mercado bancário brasileiro: 85,24% do consolidado bancário I e 84,34% de todo o Sistema Financeiro Nacional (SFN)⁷.

3. Modelagem do Sistema de Avaliação Fuzzy

3.1. SISTEMA DE INFERÊNCIA FUZZY - SIF

A construção de um modelo de avaliação é complexa e controversa. Mesmo havendo consenso no que se trata da definição de quais indicadores a acompanhar, analistas diferentes podem com o mesmo leque de informações obterem conclusões divergentes. É exatamente neste ponto, que a lógica nebulosa pretende ser útil para construção de um modelo de avaliação nas quais as incertezas e imprecisões presentes na análise de desempenho corporativo sejam consideradas.

De acordo com Matarazzo (1997), três são tipos básicos de avaliações de um índice: pelo significado intrínseco, pela comparação ao longo de vários exercícios e pela comparação com padrões. A análise do valor intrínseco de um índice é limitada e só deve ser utilizada quando não se dispõe de índices-padrão proporcionados pela análise de um conjunto de empresas. A análise pela comparação ao longo de vários exercícios revela-se útil apenas para mostrar tendências seguidas pela empresa. Assim, para atingimento do objetivo do trabalho analisam-se os índices comparando-os a padrões de referência do sistema bancário nacional.

A avaliação de um índice e a sua conceituação como bom, médio ou ruim só pode ser feita através da comparação com padrões. Não existe o bom ou o ruim em sentido absoluto, o bom só é bom em relação a outros elementos. Assim, é preciso definir um conjunto (universo) e, em seguida comparar um elemento específico com os demais elementos do conjunto para atribuir-lhe determinada qualificação. Esse é um processo natural do raciocínio humano (Matarazzo, 1997). O sistema proposto procura emular o processo de avaliação dos bancos segundo os indicadores escolhidos. Para tal, emprega-se o *software* Matlab® (*Matrix Laboratory*).

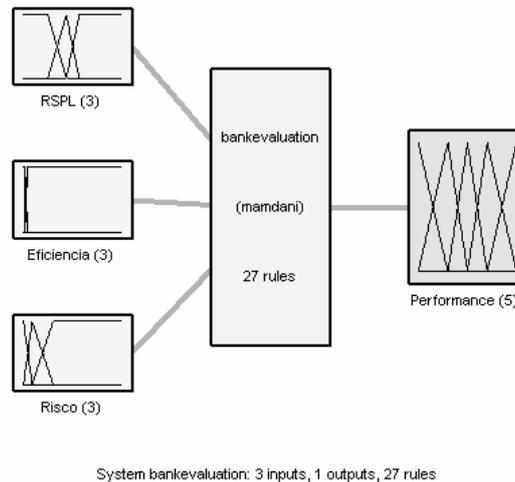
O primeiro modelo sugestionado é baseado nos conceitos de lógica nebulosa. A construção da árvore de decisões, ou seja, a determinação das variáveis lingüísticas de entrada/saída, a escala de valores para cada variável, as regras de conduta e o tipo de

⁶ O sentido aqui empregado é: dar a conhecer os riscos, benefícios, desconfortos e implicações econômicas de procedimentos com o objetivo de que as pessoas tomem decisões devidamente esclarecidas, dentro do processo de consentimento informado.

⁷ Refere-se ao ativo bancário.

implicação e inferência, além dos operadores e tipo de função de pertinência fundamenta-se na escolha do modelo mais simples e à medida que surgem necessidades, elementos incrementais são incorporados. Em síntese, o modelo escolhido é apresentado na figura 1.

Figura 1. SIF pelo Método de Mamdani



3.2. FUZZIFICAÇÃO E FUNÇÕES DE PERTINÊNCIA

O sistema Mamdani é escolhido por sua simplicidade e eficiência, além de bastante condizente com a intuição humana. As partições dos conjuntos de entrada em três (ruim, médio e bom) e de saída em cinco (péssimo, ruim, satisfatório, bom e ótimo) são recomendadas por facilitarem o processo de criação de regras pelos especialistas. Uma análise descritiva se fez adequada para determinação do domínio e suporte das variáveis de entrada⁸. Em Matarazzo (1997) um sistema de classificação relativa para índices é proposta via posição em termos de decis. Logo, com o ajuste de cinco para três variáveis lingüísticas para as entradas do sistema proporciona um método coerente de determinação dos conjuntos *fuzzy* (vide tabela 1). O domínio da saída é de 0 a 10, ou seja, um sistema tradicional de notas.

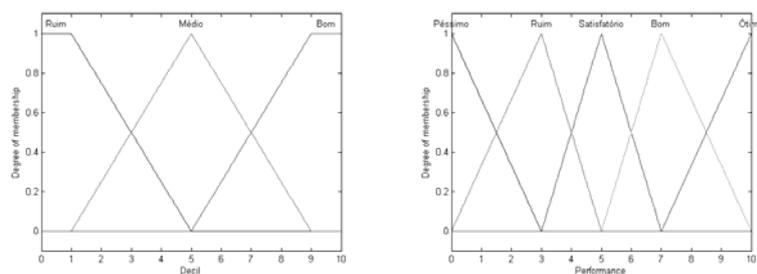
Tabela 1. Decis e classificação lingüística

		RSPL (%)	Eficiência (%)	Risco de Crédito (%)	Classificação Matarazzo	Classificação Ajustada
Pior Índice		-47,55	2340,36	34,26		
Decil	1	-11,48	101,62	15,12	Péssimo	Ruim
	2	1,63	77,72	9,79	Deficiente	
	3	9,51	67,47	7,73	Fraco	
	4	12,07	58,90	6,49	Razoável	Satisfatório
	5	15,67	53,28	4,53	Satisfatório	
	6	17,92	49,73	3,32		Bom
	7	22,67	44,66	2,51		
	8	27,29	39,55	1,57	Ótimo	
	9	34,37	28,55	0,20		
Melhor Índice		97,61	1,96	0,00		Bom

⁸ Para análise estatística descritiva usou-se o software SPSS® (*Statistical Package for the Social Sciences*).

Os formatos das funções de pertinência são trapezoidais nos extremos e triangulares no meio, pressupõe-se que para a maioria das aplicações do mundo real, a aproximação linear é suficiente. O quinto decil é o valor típico para o termo linguístico médio, e a função de pertinência decresce até atingir o mínimo nos decis 1 e 9. Nestes dois últimos, a função de pertinência dos conjuntos extremos (bom ou ruim) atinge o máximo (um) reduzindo o efeito de *outliers*⁹. Para a saída, os conjuntos são triangulares, onde cada conjunto é representado com grau de pertinência total pelo valor mais representativo (figura 2).

Figura 2. Função de Pertinência Geral das Variáveis de Entrada e Função de Pertinência



3.3. REGRAS DE AVALIAÇÃO

O próximo passo é a definição das regras de avaliação. Estas, também conhecidas como inferência *fuzzy*, aplicam as regras às entradas *fuzzy* geradas no processo de *fuzzificação*. A regra *min-max* é a utilizada, onde considera o mínimo dos antecedentes para determinar as regras de maior força e o máximo das regras de maior força para cada conseqüente, originando assim as saídas *fuzzy*. Em resumo: utiliza-se o mecanismo de inferência conhecido como *modus ponens generalizado*, com o mínimo para o conectivo *e*, mínimo para a implicação e máximo para agregação.

O particionamento de entrada gera um número máximo de 27 regras possíveis. Estas foram obtidas por pesquisa junto a sete especialistas da área de avaliação de desempenho corporativo¹⁰.

3.4. DEFUZZIFICAÇÃO

O módulo de defuzzificação traduz o estado da variável *fuzzy* de saída para um valor numérico, no caso a nota final do modelo de avaliação que mensura a performance do banco relativo aos demais concorrentes e ao período específico de análise.

São cinco os métodos automáticos disponibilizados pelo Matlab®, a saber: *Centroid*, *Bisector*, *MOM*, *LOM*, *SOM*. Não existe consenso na literatura a respeito do melhor método. Para Shaw e Simões (2001), o MOM é o mais apropriado para decisões qualitativas como avaliação de crédito, já Von Altrock (1997) argumenta que o *centroid* é preferível no campo das ciências sociais aplicadas. Para escolha do modelo de avaliação mais adequado compara-se cada método com as respostas obtidas por pesquisa junto aos especialistas.

Para isto, o total de 97 bancos é dividido em dois grupos por tamanho do ativo: os 49 maiores e os 48 menores. De cada grupo selecionam-se aleatoriamente duas amostras de

⁹ Os conjuntos extremos, ruim ou bom, são invertidos para o caso dos indicadores quanto menor melhor (eficiência e risco).

¹⁰ Além da experiência de trabalho, dois especialistas são doutores, quatro são mestres e um possui especialização. As áreas de conhecimento são controladoria, economia, administração e contabilidade.

cinco bancos, perfazendo um total de vinte bancos (20,62%) divididos em quatro amostras de cinco. Dos sete especialistas consultados para a geração de regras são escolhidos aleatoriamente quatro. Para cada um é encaminhada uma amostra diferente contendo cinco conjuntos de indicadores. Os especialistas têm a opção de atribuir nota quantitativa de zero a dez para a performance de cada banco de sua respectiva amostra (os nomes de cada instituição são retirados para não influenciar a análise).

A tabela 2 apresenta as notas dos cinco métodos de defuzzificação comparados com a nota dos especialistas. O erro médio (diferença absoluta) é maior para o método LOM e menor para o MOM e *bisector*. Entretanto, para definição de quais metodologias de defuzzificação são mais apropriadas para o SIF específico (que visa gerar o melhor modelo para avaliar a performance dos bancos e, conseqüentemente, substituir o analista de avaliação corporativa) adota-se um teste de hipótese de diferenças nas médias. Cada método é comparado ao par com a série de notas dos especialistas e verifica-se a existência ou não de igualdade na média. As amostras são dependentes (dados emparelhados) e o número de pares é vinte (amostra pequena), portanto aplica-se o teste *t-student* bicaudal. Conforme observado na tabela 3, os testes para verificação da hipótese de que as diferenças entre as médias sejam nulas estatisticamente (ao nível de confiança de 95%) são rejeitadas para os métodos LOM e SOM. Conclui-se então que os três primeiros métodos (*centroid*, *bisector* e *MOM*) podem ser empregados para o intento.

Tabela 2. SIF versus Especialistas.

Empresas	Notas Especialistas	Notas Sistema de Inferência Fuzzy					Diferenças				
		Métodos					Métodos				
		Centróide	Bisector	MOM	LOM	SOM	Centróide	MOM	Bisector	LOM	SOM
BCO JOHN DEERE	7,00	6,30	6,80	7,20	8,00	6,40	0,70	-0,20	0,20	-1,00	0,60
BRADESCO	9,00	7,79	8,40	9,60	10,00	9,20	1,21	-0,60	0,60	-1,00	-0,20
PINE	9,00	8,93	9,00	9,50	10,00	9,00	0,07	-0,50	0,00	-1,00	0,00
RURAL	1,00	1,21	1,10	0,70	1,40	0,00	-0,21	0,30	-0,10	-0,40	1,00
SANTANDER BANESPA	5,00	6,12	5,60	5,00	5,50	4,50	-1,12	0,00	-0,60	-0,50	0,50
ARBI	1,00	1,21	0,90	0,10	0,20	0,00	-0,21	0,90	0,10	0,80	1,00
INDUSVAL	5,00	5,93	5,60	5,00	5,60	4,40	-0,93	0,00	-0,60	-0,60	0,60
LUSO BRASILEIRO	2,00	3,10	3,00	0,65	1,30	0,00	-1,10	1,35	-1,00	0,70	2,00
TOKYOMITSUBISHI	1,00	1,05	0,90	0,30	0,60	0,00	-0,05	0,70	0,10	0,40	1,00
TRIANGULO	9,00	7,49	7,80	9,45	10,00	8,90	1,51	-0,45	1,20	-1,00	0,10
BNP PARIBAS	7,00	7,43	7,40	7,15	8,00	6,30	-0,43	-0,15	-0,40	-1,00	0,70
CLASSICO	7,00	6,75	6,90	7,15	7,80	6,50	0,25	-0,15	0,10	-0,80	0,50
CRUZEIRO DO SUL	5,00	5,99	5,60	5,00	5,50	4,50	-0,99	0,00	-0,60	-0,50	0,50
NOSSA CAIXA	9,00	7,85	7,90	9,35	10,00	8,70	1,15	-0,35	1,10	-1,00	0,30
VOTORANTIM	7,00	7,55	7,50	7,20	8,10	6,30	-0,55	-0,20	-0,50	-1,10	0,70
BEPi	7,00	6,53	6,70	7,20	8,20	6,20	0,47	-0,20	0,30	-1,20	0,80
BVA	1,00	1,34	1,10	0,55	1,10	0,00	-0,34	0,45	-0,10	-0,10	1,00
GUANABARA	6,00	5,93	6,60	7,15	7,70	6,60	0,07	-1,15	-0,60	-1,70	-0,60
LEMON BANK	2,00	3,21	3,50	0,65	1,30	0,00	-1,21	1,35	-1,50	0,70	2,00
PARANA	9,00	9,02	9,10	9,85	10,00	9,70	-0,02	-0,85	-0,10	-1,00	-0,70
Média	5,45	5,54	5,57	5,44	6,02	4,86	-0,09	0,01	-0,12	-0,57	0,59
Desvio Padrão	3,07	2,67	2,81	3,63	3,67	3,60	0,80	0,66	0,65	0,71	0,69
Valor Máximo	9,00	9,02	9,10	9,85	10,00	9,70	1,51	1,35	1,20	0,80	2,00
Valor Mínimo	1,00	1,05	0,90	0,10	0,20	0,00	-1,21	-1,15	-1,50	-1,70	-0,70
Média Absoluta*							0,63	0,49	0,49	0,83	0,74
Desvio Padrão Absoluto**							0,48	0,43	0,43	0,35	0,52

*Média da série de diferenças absolutas.

**Desvio padrão da série de diferenças absolutas.

Tabela 3. Teste de hipóteses para os Sistemas de Inferência Fuzzy.

	Média	Variância	Observações	Correlação de Pearson	Hipótese da diferença de média	GL	Estatística t	t crítico bi-caudal	P(T<=t) bi-caudal
Especialistas	5,45	9,42	20						
Centróide	5,54	7,15	20	0,97	0	19	0,49	2,09	0,63
Bisector	5,57	7,88	20	0,98	0	19	0,83	2,09	0,42
MOM	5,44	13,18	20	0,99	0	19	-0,08	2,09	0,93
LOM	6,02	13,48	20	0,99	0	19	3,55	2,09	0,00
SOM	4,86	12,98	20	0,99	0	19	-3,81	2,09	0,00

4. Modelagem do Sistema de Avaliação Neuro-Fuzzy

Apesar de demonstrar a capacidade de emular o modelo mental de avaliação dos

especialistas consultados, o sistema elaborado na seção anterior, ainda possui algumas limitações.

Apesar de alguns especialistas estarem acostumados a trabalhar com a escala original proposta em Matarazzo (1997), onde sete variáveis linguísticas são empregadas para representar o desempenho relativo de um determinado indicador, no modelo implementado anteriormente apenas três partições são efetuadas para as variáveis de entrada. Isto por motivos de simplificação, ou mesmo viabilização na criação das regras pelos especialistas. Caso optasse pelo particionamento das três variáveis de entrada em sete e não em três conjuntos *fuzzy*, o total de regras possíveis seria 343 e não as 27 originais. Para os especialistas a criação de 343 regras se mostra inadequada e bastante suscetível a erros.

Portanto, paralelamente ao *SIF* sugerido testa-se também um modelo híbrido de avaliação da performance bancária que alia a lógica *fuzzy* às redes neurais: o modelo *neuro-fuzzy*. A capacidade de aprendizado, generalização e extração de regras pelas redes neurais são úteis para casos em que as regras não estão disponíveis, e/ou não se tem capacidade de definição do particionamento das variáveis de entrada.

O modelo específico utilizado é o *ANFIS* (*Adaptive Neuro-Fuzzy Inference System*), também disponibilizado pelo Matlab®. Este modelo é capaz, por meio dos dados de entrada e saída, construir um *SIF*, no qual as funções de pertinência são ajustadas pelo algoritmo *backpropagation*, ou pela combinação deste com o método *least squares* (híbrido). Permite-se assim, que o sistema *fuzzy* aprenda com os dados e molde sua estrutura com menor interferência do analista. O *ANFIS* é um sistema de Takagi-Sugeno e demonstra-se que seu modelo de primeira ordem é um aproximador universal¹¹. O tipo de particionamento que o sistema utiliza é *fuzzy-grid* adaptativo, no qual seu aprendizado (identificação da estrutura e parâmetros) é feito em duas etapas que se repetem até que o critério de parada seja alcançado. A popularidade do *ANFIS* é tão grande que o levou a ser implementado no Matlab® (Morais *et al.*, 2005).

Para utilização do *ANFIS* é necessário um pré-processamento específico dos dados. Primeiro, leva-se em consideração as informações descritivas das variáveis de entrada, sendo a normalização efetuada em partes. Os valores identificados pelo SPSS® como *outliers* são representados por 0 ou 1 (*outlier* inferior ou superior, respectivamente), os demais são normalizados entre 0 e 1. Para aprender com os dados é necessário o fornecimento dos pares entrada/saída. Desta forma, uma nova solicitação de notas para os bancos restantes (77 não avaliados na primeira etapa) foi solicitada a quatro especialistas¹².

Várias tentativas de modelos são geradas buscando sempre a parcimônia, deste modo, o *trade-off* convergência e generalização é fator preponderante na escolha das melhores configurações. O *ANFIS* permite que seja realizada a validação cruzada pela opção *checking data set*, minimizando a possibilidade de ocorrência de *overfitting*. Assim sendo, a amostra é dividida aleatoriamente nas seguintes proporções aproximadas: 70 % (68 bancos) para o conjunto de treinamento, 20% (19 bancos) para o conjunto de validação e 10% (10 bancos) para o conjunto de teste.

Além da configuração dos conjuntos de treinamento, validação e teste nas proporções acima, para todos os modelos testados utilizam-se: tipo de particionamento *grid partition*; método de otimização híbrido; operador de produto para representar o conectivo *e*; média ponderada para a defuzzificação; zero de tolerância ao erro no critério de parada (não se sabe *a priori* qual o comportamento do erro). O número de épocas define-se pela análise dos erros nos conjuntos de treinamento e validação. O crescimento contínuo do erro no conjunto de

¹¹ O Matlab® permite modelos do tipo Sugeno de ordem zero ou um.

¹² Três analistas avaliaram 19 bancos cada e um avaliou 20.

validação determina a parada do treinamento. A partir do número obtido os parâmetros da rede são “zerados” e um novo treinamento é gerado baseando-se no número de épocas escolhido na primeira etapa¹³.

The main difference between Mamdani and Sugeno is that the Sugeno output membership functions are either linear or constant (THE MATHWORKS, 2002). Logo, para fins de comparação às primeiras tentativas utilizam funções de pertinência iguais às utilizadas no *SIF* pelo método de Mamdani proposto anteriormente (formato e número de conjuntos) sem a utilização da rede neural para ajuste das funções de pertinência. No geral, os resultados não são satisfatórios. O modelo de melhor aderência ao conjunto de teste (menor erro médio) tem o aproximador universal do tipo Takagi-Sugeno de primeira ordem, é treinado em 16 épocas e tem os seguintes níveis de erro médio: 0.12 (treinamento), 4.08 (validação), e 4.73 (teste).

No segundo grupo de tentativas permite-se a definição dos conjuntos pelo *ANFIS*¹⁴, sendo testadas todas as possíveis combinações de modelos variando: o formato das funções de pertinência (triangular, trapezoidal ou gaussiana), a quantidade de conjuntos *fuzzy*¹⁵ (três, cinco ou sete), e a função de saída (constante ou linear). Os modelos com menor erro médio no conjunto de teste estão na tabela 4.

Tabela 4. Modelos selecionados da amostra original.

	Partição	F.Pertinência	Saída	Épocas	Erro		
					Treinamento	Validação	Teste
Modelo 1	3	Triangular	Constante	5	0,74	0,96	4,17
Modelo 2	3	Trapézio	Constante	14	0,75	0,73	2,17
Modelo 3	3	Gaussiana	Linear	15	0,12	4,08	4,72
Modelo 4	5	Triangular	Constante	8	0,22	4,45	10,52

Os modelos acima apresentam erros (média) no conjunto de teste superiores aos encontrados no *SIF* pelo método de Mamdani (por qualquer das três defuzzificações aceitas). No entanto, um problema relevante identificado na geração dos modelos acima é o tamanho da amostra de pares entrada/saída (97 elementos). O número é pequeno para a geração de regras e otimização do modelo dado o número de parâmetros necessários à rede. Este problema ocasiona resultados ruins para os sistemas no qual o particionamento seja superior a três (como exemplo o modelo 4).

Para contornar o problema da insuficiência de dados visando ao adequado ajuste dos parâmetros, duplica-se a amostra. A constituição dos conjuntos de treinamento, validação e teste se mantiveram nas mesmas proporções (duplica-se o valor absoluto de cada conjunto), sendo porém, os bancos distribuídos de forma aleatória novamente. Todas as combinações possíveis de modelos foram testadas mais uma vez. A tabela 5 condensa os modelos com menor erro médio no conjunto de teste.

¹³ Um exemplo de *SIF* pelo método Sugeno utilizado no artigo encontra-se na figura A.1 (ver apêndice).

¹⁴ O modelo geral utilizado no trabalho está representado na figura A.2.

¹⁵ Sempre a mesma para todas as três entradas.

Tabela 5. Modelos selecionados da amostra duplicada.

	Partição	F.Pertinência	Saída	Épocas	Erro		
					Treinamento	Validação	Teste
Modelo 5	3	Triangular	Constante	8	0,86	0,85	0,96
Modelo 6	3	Trapézio	Constante	14	0,84	0,98	1,16
Modelo 7	3	Gaussiana	Constante	80	0,67	0,75	0,75
Modelo 8	3	Gaussiana	Linear	25	0,09	0,75	0,49
Modelo 9	5	Gaussiana	Constante	14	0,14	1,93	1,02
Modelo 10	5	Gaussiana	Linear	10	0,001	0,48	0,29
Modelo 11	7	Gaussiana	Constante	8	0,002	1,69	1,65

O aumento do número de observações melhora o processo de aprendizado da rede reduzindo o erro médio no conjunto de teste (a comparação do modelo 1 com o modelo 5, do 2 com o 6 e do 3 com o 8, são exemplos). Possibilita também, encontrar modelos com particionamento de cinco ou sete, e conseqüentemente maior número de regras, que tenham erros médio semelhantes ou menores aos encontrados no particionamento de três. Inclusive, o sistema com menor erro médio (0,29) no conjunto de teste para a amostra duplicada é o modelo 10, com cinco conjuntos *fuzzy* para cada variável.

O teste de hipóteses de igualdade de média é realizado para os sistemas neuro-*fuzzy* com menor erro médio, tanto os gerados pela amostra original quanto pela amostra duplicada (modelo 2 e 10, respectivamente). Como o conjunto de testes é diferente para cada amostra, inclusive quanto ao tamanho, do conjunto de testes duplicado retira-se a metade (dez bancos) randomicamente conforme tabela 6.

Tabela 6. Teste de hipóteses para os Sistemas de Inferência Neuro-*Fuzzy*.

Conj. Teste A. Original	Especialistas	Modelo 2	Conj. Teste A. Duplicada*	Especialistas	Modelo 10
BNB	5,00	5,01	BB	8,00	7,99
CLASSICO	3,00	3,00	FATOR	9,00	9,00
BCO JOHN DEERE	3,00	3,03	SOCOPA	5,00	5,90
TRIANGULO	7,00	6,97	SAFRA	7,00	7,00
LA PROVINCIA	5,00	5,00	BANIF	2,00	2,00
BANKBOSTON	1,00	1,01	RURAL	0,00	1,00
INTERCAP	7,00	7,00	SANTANDER BANESPA	6,00	6,00
BEPI	4,00	4,03	TRIANGULO	7,00	7,00
POTTENCIAL	5,00	5,00	INDUSTRIAL DO BRASIL	7,00	6,99
PROSPER	5,00	5,12	BRADESCO	9,00	9,00
Média	4,50	4,52	Média	6,00	6,19
Desvio Padrão	1,84	1,84	Desvio Padrão	2,94	2,70
Variância	3,39	3,37	Variância	8,67	7,29
Observações	10	10	n	10	10
		1,00	Correlação de Pearson		0,99
Hipótese da diferença de média		0,00	Hipótese da diferença de média		0,00
GL		9,00	GL		9,00
Estatística t		-1,41	Estatística t		-1,48
t crítico bi-caudal		2,26	t crítico bi-caudal		2,26
P(T<=t) bi-caudal		0,19	P(T<=t) bi-caudal		0,17

*Sub-amostra de 10 bancos tomados aleatoriamente.

Observa-se que os dois modelos escolhidos não são rejeitados ao nível de significância estatística de 5%. Portanto, são também passíveis de utilização para auxílio aos

especialistas quanto à avaliação corporativa de bancos e a geração de *rankings* de desempenho¹⁶.

5. Conclusão

A lógica nebulosa permite tratar, de forma numérica, características como péssimo, ruim, satisfatório, bom, ótimos, entre outros. Como decorrência dos adequados tratamentos, o resultado da avaliação é expresso em um valor numérico que representa a avaliação quantitativa do desempenho dos bancos. A possibilidade de lidar com informações imprecisas torna-se “trunfo” de um modelo de avaliação *fuzzy*. As ambigüidades inerentes ao processo avaliatório podem ser tratadas de forma quantitativa em um modelo deste tipo.

A capacidade de aprendizado de uma rede neural aliada ao aproveitamento do conhecimento humano de um sistema *fuzzy* pode contribuir para sistemas de avaliação onde o analista não seja capaz de explicitar as regras ou o melhor formato das funções de pertinência.

Os modelos propostos podem auxiliar analistas de análise corporativa de bancos de várias maneiras. Grandes diferenças na avaliação de um determinado analista comparadas ao do modelo podem ser identificadas e verificadas *a posteriori* com maior cuidado. Analistas com pouca experiência passam a ter uma ferramenta de parametrização e que corrobora para aceleração do aprendizado.

A capacidade de pontuação da performance dos bancos segundo indicadores-padrão baseados não só na rentabilidade, mas na eficiência e risco, em determinado período, possibilita a criação de um *ranking* de desempenho bancário multidimensional.

Para o artigo em questão, tanto o sistema de inferência *fuzzy* tradicional, quanto o neuro-*fuzzy* mostram-se capazes de emular o modelo mental dos analistas pesquisados. Faz-se necessário, no entanto, salientar algumas limitações do trabalho. Para um *ranking* nacional de bancos um número maior de especialistas deve ser consultado, bem como uma amostra maior (avaliações trimestrais e de períodos anteriores) deve ser utilizada, principalmente para o sistema híbrido. Novos índices podem ser empregados, por exemplo, indicadores que meçam a relação entre bancos e clientes seriam úteis para uma visão de futuro da instituição. Portanto, a pesquisa empreendida não pretende ser conclusiva, mas sim demonstrar a viabilidade da aplicação de técnicas inteligentes para o processo avaliativo.

Referências

Carvalho, F. J. C. Inovação financeira e regulação prudencial: da regulação de liquidez aos acordos da Basiléia. In: Rogério Sobreira. (Org.). *Regulação Financeira e Bancária*. 1º ed. São Paulo: Atlas, v. 1, p. 121-137, 2005.

Schwartzman, S. *A Problemática da Avaliação: excelência acadêmica e maturação institucional*, Brasília, CEDATE, 1986. Disponível em: http://www.schwartzman.org.br/simon/p_aval.htm

Miranda, L. C.; Silva, J. D. G. *Medição de desempenho. Controladoria: agregando valor para a empresa*. Porto Alegre, Bookman, 2002.

¹⁶ No apêndice (tabela A.I) estão disponibilizadas os *rankings* criados pelo SIF com defuzzificação MOM e pelo modelo 10 (neuro-*fuzzy*). Apesar de algumas diferenças na classificação dos bancos os dois modelos agrupam (por posição) de foram semelhante.

Falkenstein, E. Accounting for Economic and Regulatory Capital in RAROC Analysis. *Bank Accounting and Finance*, v. 11, nº 11, fall 1997.

James, C. RAROC Based Capital Budgeting and Performance Evaluation: A Case Study of Bank Capital Allocation, University of Pannsylvania, September 1996. Disponível em: <http://fic.wharton.upenn.edu/fic/papers/96/9640.pdf>

Matarazzo, D. C. *Análise Financeira de Balanços: abordagem básica e gerencial*. 4º ed. São Paulo: Atlas, 1997.

Banco Central do Brasil. 50 Maiores Bancos e o Consolidado do Sistema Financeiro Nacional. Disponível em: <http://www.bcb.gov.br/fis/top50/port/default.asp?parmidiaoma=P&id=top50>. Consultado em 27/06/2006.

Shaw, I. S. e Simões, M. G. *Controle e Modelagem Fuzzy*. São Paulo: Edgard Blücher, 2001.
Von Altrock, C. *Fuzzy Logic and NeuroFuzzy applications in business and finance*. New Jersey: Prentice Hall, 1997.

Morais Júnior, A. M. F.; Holanda, U.; Souza, A. M.; Tostes, M. E. L. e Costa Júnior, C. T. Sistema Inteligente para Previsão de Carga em Transformadores de Distribuição de Energia: estudos na frequência fundamental e em frequências harmônicas. *VI SBQEE*, Belém, agosto 2005.

The Mathworks, INC. *Fuzzy Logic Toolbox User's Guide*, copyright 1995–2002.

Apêndice

Figura A.1. SIF pelo Método Sugeno: um exemplo

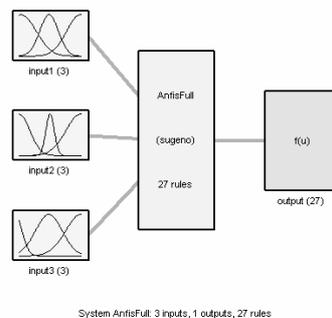


Figura A.2. Estrutura Geral ANFIS: 5 Camadas e 27 Regras

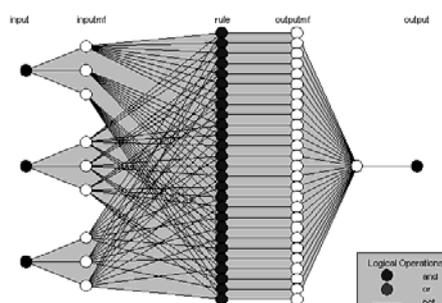


Tabela A.1. Teste de hipóteses para os Sistemas de Inferência Neuro-Fuzzy

Ranking	Bancos	Modelo 10	Ranking	Bancos	SIF Mamdani MOM	Ranking	Bancos	Modelo 10	Ranking	Bancos	SIF Mamdani MOM
1	DEUTSCHE	10,00	1	DEUTSCHE	10,00	50	CACIQUE	5,00	50	BIC	5,00
2	PACTUAL	9,23	2	PARANA	9,85	51	BASA	5,00	51	BASA	5,00
3	BRADESCO	9,00	3	FATOR	9,75	52	CRUZEIRO DO SUL	5,00	52	MERCANTIL DO BRASIL	5,00
4	BONSUCESSO	9,00	4	BONSUCESSO	9,70	53	CR2	5,00	53	FIBRA	5,00
5	NOSSA CAIXA	9,00	5	BRADESCO	9,60	54	CITIBANK	5,00	54	SS	5,00
6	PINE	9,00	6	HSBC	9,50	55	MERCANTIL DO BRASIL	5,00	55	CRUZEIRO DO SUL	5,00
7	BMG	9,00	7	PINE	9,50	56	CREDIBEL	5,00	56	BRB	5,00
8	GERDAU	9,00	8	BANESE	9,50	57	MAXIMA	5,00	57	BEC	5,00
9	ITAU	9,00	9	TRIANGULO	9,45	58	LA PROVINCIA	5,00	58	SOFISA	5,00
10	BANESTES	9,00	10	GERDAU	9,45	59	POTTENCIAL	5,00	59	BGN	5,00
11	FATOR	9,00	11	BB	9,40	60	KEB	5,00	60	SCHAHIN	5,00
12	PARANA	9,00	12	BMG	9,40	61	BANCAP	5,00	61	BANPARA	5,00
13	BBM	9,00	13	NOSSA CAIXA	9,35	62	FIBRA	5,00	62	CACIQUE	5,00
14	HSBC	8,18	14	PACTUAL	9,35	63	SS	5,00	63	INDUSVAL	5,00
15	RENDIMENTO	8,00	15	RENDIMENTO	9,35	64	BRB	5,00	64	MODAL	5,00
16	BANESE	8,00	16	ITAU	9,30	65	BNB	5,00	65	PROSPER	5,00
17	VOTORANTIM	8,00	17	BBM	9,30	66	RENNER	5,00	66	SOCOPA	5,00
18	BB	7,99	18	BANESTES	9,30	67	ABN AMRO	5,00	67	CREDIBEL	5,00
19	DAYCOVAL	7,00	19	CEF	7,25	68	PROSPER	5,00	68	RENNER	5,00
20	ABC-BRASIL	7,00	20	SAFRA	7,25	69	CEDULA	5,00	69	CEDULA	5,00
21	BMC	7,00	21	JP MORGAN CHASE	7,25	70	INDUSVAL	4,00	70	MAXIMA	5,00
22	JP MORGAN CHASE	7,00	22	CREDIT SUISSE	7,25	71	BEPI	4,00	71	LA PROVINCIA	5,00
23	BANRISUL	7,00	23	INDUSTRIAL DO BRASIL	7,25	72	BANPARA	4,00	72	POTTENCIAL	5,00
24	CREDIT SUISSE	7,00	24	VOTORANTIM	7,20	73	LA REPUBLICA	4,00	73	KEB	5,00
25	IBIBANK	7,00	25	ALFA	7,20	74	GUANABARA	3,00	74	CR2	5,00
26	BANCOOB	7,00	26	ABC-BRASIL	7,20	75	CLASSICO	3,00	75	LA REPUBLICA	5,00
27	VR	7,00	27	IBIBANK	7,20	76	BCO JOHN DEERE	3,00	76	BANCAP	5,00
28	INTERCAP	7,00	28	BANCOOB	7,20	77	MORADA	3,00	77	MORADA	2,90
29	MATONE	7,00	29	DAYCOVAL	7,20	78	LUSO BRASILEIRO	2,25	78	RURAL	0,70
30	BANSICREDI	7,00	30	BCO JOHN DEERE	7,20	79	LEMON BANK	2,23	79	BANIF	0,70
31	PECUNIA	7,00	31	VR	7,20	80	BANIF	2,00	80	FICSA	0,70
32	CEF	7,00	32	PECUNIA	7,20	81	RABOBANK	2,00	81	EMBLEMA	0,70
33	RIBEIRAO PRETO	7,00	33	BEPI	7,20	82	EMBLEMA	2,00	82	RABOBANK	0,65
34	SAFRA	7,00	34	MATONE	7,20	83	CREDIT LYONNAIS	1,35	83	AMEX	0,65
35	TRIANGULO	7,00	35	BNP PARIBAS	7,15	84	BANKBOSTON	1,00	84	LUSO BRASILEIRO	0,65
36	ALFA	7,00	36	CLASSICO	7,15	85	RURAL	1,00	85	LEMON BANK	0,65
37	INDUSTRIAL DO BRASIL	6,99	37	BANSICREDI	7,15	86	AMEX	1,00	86	BVA	0,55
38	BRASCAN	6,82	38	INTERCAP	7,15	87	TOKYOMITSUBISHI	1,00	87	BANCNACION	0,55
39	BIC	6,00	39	GUANABARA	7,15	88	GE CAPITAL	1,00	88	BANKBOSTON	0,50
40	SCHAHIN	6,00	40	BMC	7,10	89	BVA	1,00	89	GE CAPITAL	0,40
41	SOFISA	6,00	41	BRASCAN	7,10	90	FICSA	1,00	90	TOKYOMITSUBISHI	0,30
42	SANTANDER BANESPA	6,00	42	RIBEIRAO PRETO	7,10	91	BANCNACION	1,00	91	BESC	0,25
43	SOCOPA	5,90	43	BANRISUL	7,05	92	ARBI	1,00	92	ARBI	0,10
44	BARCLAYS	5,25	44	BARCLAYS	7,05	93	BESC	0,34	93	CREDIT LYONNAIS	0,05
45	MODAL	5,00	45	UNIBANCO	5,00	94	SOCIETE GENERALE	0,00	94	ING	0,00
46	UNIBANCO	5,00	46	SANTANDER BANESPA	5,00	95	BCOMURUGUAI	0,00	95	SOCIETE GENERALE	0,00
47	BGN	5,00	47	ABN AMRO	5,00	96	ING	0,00	96	BCOMURUGUAI	0,00
48	BEC	5,00	48	CITIBANK	5,00	97	CARGILL	0,00	97	CARGILL	0,00
49	BNP PARIBAS	5,00	49	BNB	5,00						