APLICAÇÃO DE REDES NEURAIS NA CLASSIFICAÇÃO DE RENTABILIDADE FUTURA DE EMPRESAS

Élia Yathie Matsumoto[1] e Afonso de Campos Pinto[2]

[1]Opencadd Advanced Technology – Av. Brig. Faria Lima, 1931 – 15º Andar - Cj 152 – São Paulo/SP/01452-001
[2]Fundação Getúlio Vargas – Escola de Economia de São Paulo - Rua Itapeva, 474 – São Paulo/SP/01332-00

elia@opencadd.com.br ; afonso.pinto@fgv.br

Resumo – Este trabalho objetiva mostrar a aplicação de Redes Neurais para estimar a classificação de rentabilidade de empresas e, desta forma, prover suporte para o desenvolvimento de sistemas de apoio à tomada de decisão de investimentos. Para tanto, a partir de dados financeiros e contábeis das 1000 melhores empresas listadas anualmente pela publicação Melhores e Maiores – Exame (Editora Abril), entre 1996 e 2005, classificou-se a rentabilidade das empresas em três categorias: retorno ruim (investimento não recomendado), retorno médio (investimento recomendado) e retorno ótimo (investimento altamente recomendado). Para serem comparados com o modelo de Redes Neurais, foram escolhidos o modelo clássico de Regressão Linear Múltipla, como referência mínima, e o de Regressão Logística Ordenada, como marca comparativa de desempenho (*benchmark*). Neste experimento, os três modelos, construídos com as informações das empresas entre 1996 e 2005, receberam dados de 2005 para estimar a classificação das empresas em 2006. Os resultados obtidos forneceram evidências de que os modelos de Redes Neurais conseguem gerar resultados expressivamente superiores quando comparados aos modelos de Regressão Linear Múltipla e de Regressão Logística Ordenada.

Palavras-chave – Redes Neurais, Regularização Bayesiana, Classificação de Padrões, Regressão Linear Múltipla, Regressão Logística Ordenada, Rentabilidade de Empresas, Finanças Corporativas.

1. Introdução

Estudos sobre o uso de Redes Neurais (RN) em finanças podem ser encontrados em um amplo conjunto de aplicações: negociações no mercado financeiro [1], previsão de desempenho no mercado financeiro [2-4], modelamento de taxa de câmbio [5], previsão de correlação entre moedas [6], previsão de evolução de fundos mútuos [7], predição de insolvência [8], entre outras. No Brasil, muitos trabalhos têm sido desenvolvidos abordando especificamente as condições do mercado local, e.g. análise de crédito [9], apreçamento de opções no mercado brasileiro [10], [11], recuperação de receita com prevenção de insolvência [12], avaliação de crédito bancário em instituição financeira brasileira [13] e variação de cotação de ações da Telebrás [14]. Boa parte desses estudos elabora comparações entre RN com outros tipos de modelos, e.g. lineares [5], *k-means* [13] e Black & Scholes [10], [11].

Particularmente, com relação à aplicação de RN em estudos sobre previsão de rentabilidade e desempenho de empresas, foram publicados trabalhos com abordagem de classificação binária, como o de Almeida e Dumontier [15], no qual os autores, tendo como dados de entrada informações contábeis das empresas do setor de carga rodoviária francês, classificaram as mesmas como possivelmente inadimplentes ou não. O estudo indicou que o modelo de redes neurais não conseguiu produzir resultados superiores aos modelos desenvolvidos com métodos estatísticos tradicionais. Uma das explicações dada pelos autores para esse insucesso foi a limitação para explorar soluções mais sofisticadas imposta pelo excessivo esforço de desenvolvimento e tempo de processamento exigido pelos modelos de RN. Considerando-se que o experimento foi realizado há mais de treze anos, provavelmente essas barreiras já devem ter sido superadas.

Sobre previsão de desempenho, o trabalho escrito por Lam [16] descreve um modelo de RN que recebe, como dados de entrada, um conjunto de informações financeiras e contábeis de empresas, além de um conjunto de informações macroeconômicas, produzindo como resultado as taxas de retorno de suas ações para tomada de decisão de investimentos. Para verificar o desempenho do modelo de RN, os resultados dos investimentos indicados pela RN foram comparados com um portfólio de rendimento

mínimo, composto por uma cesta de empresas bem diversificada (portfólio *média do mercado*) e com um portfólio de rendimento máximo, composto pelas empresas presentes na região de 1/3 do topo da lista de taxas de retorno mais altas (portfólio *informação perfeita*). O estudo concluiu que o modelo de RN conseguiu superar o portfólio de rendimento mínimo, mas ficou abaixo do de rendimento máximo. Não foram encontradas evidências de que a introdução de variáveis macroeconômicas melhora o modelo de RN.

Seguindo a linha de previsão de desempenho futuro de empresas, este trabalho também se utilizou de informações financeiras e contábeis como dados de entrada para os modelos de RN, de forma semelhante aos dois estudos mencionados acima. Porém, não se tentou prever um valor numérico, como no caso do trabalho de Lam [16]. Foi proposta uma classificação da rentabilidade futura das empresas analisadas em três classes (rentabilidade *ruim*, *média* e *ótima*), ao invés de duas como no caso do estudo de Almeida e Dumontier [15]. Para efeito de comparação com o modelo proposto de RN, foram escolhidos (*i*) o modelo clássico de Regressão Linear Múltipla como referência mínima, e (*ii*) o de Regressão Logística Ordenada como marca comparativa de desempenho (*benchmark*). A metodologia e os parâmetros aqui adotados são detalhados no item a seguir. Resultados numéricos são apresentados no item 3. Uma discussão sobre os possíveis efeitos de variações nas variáveis de entrada sobre as variáveis de saída é explorada no item 4. As principais conclusões do trabalho são apresentadas no item 5.

2. Metodologia

Neste trabalho utilizamos a metodologia descrita por Duda, Stork e Hart [17] para a solução de problemas de classificação de padrões, na qual o processo é organizado em seis passos:

(I) Definição do conjunto de indivíduos;
(II) Identificação dos atributos dos indivíduos;
(III) Definição de um subconjunto de dados de treinamento e outro de verificação;
(IV) Definição dos modelos de estimação;
(V) Construção dos modelos utilizando o subconjunto de dados de treinamento;
(VI) Avaliação dos modelos por meio da comparação dos resultados obtidos com o subconjunto de dados de verificação.

Com relação ao Passo IV, como já mencionado, foram escolhidos o modelo clássico de Regressão Linear Múltipla (RLM) e o de Regressão Logística Ordenada (RLO) para serem comparados com o modelo de RN. Os demais passos mencionados nortearam a preparação dos dados e os experimentos descritos a seguir. Todos os modelos foram desenvolvidos com o software MATLAB 7.6 com os módulos *Statistics Toolbox* e *Neural Network Toolbox*.

2.1. Descrição e Preparação dos Dados

O problema proposto consistiu em estimar a classificação da rentabilidade de uma empresa no ano seguinte, dado um conjunto de informações financeiras do ano corrente, utilizando modelos treinados com informações passadas. Para tanto, utilizamos os dados financeiros e contábeis de empresas classificadas na pesquisa '1000 Melhores & Maiores', no período entre 1996 e 2006 [18]. Foram selecionadas empresas estatais e do setor privado e, excluídas instituições financeiras, bancos e seguradoras.

Dos 24 parâmetros disponíveis, 11 foram selecionados para a construção dos modelos, a saber:

1. Capital Circulante Líquido (*CCL*);
2. Lucro Líquido Ajustado (*LLA*);
3. Patrimônio Líquido Ajustado (*PLA*);
4. Riqueza Criada por Empregado (*RCE*);
5. Crescimento das Vendas (*CV*);
6. Endividamento Geral (*EG*);
7. Margem das Vendas (*MV*);
8. Índice de Giro de Ativo (*IGA*);
9. Índice de Liquidez Geral (*ILG*);
10. Vendas em Dólares (*VD*);
11. Valor da Rentabilidade do Patrimônio Ajustado (*VR*).

A partir desse conjunto de parâmetros, criamos as observações da seguinte forma:

- Variável dependente: a partir do parâmetro VR_t, Valor da Rentabilidade do Patrimônio Ajustado em t, dos anos de 1997 a 2006, construímos a variável dependente do modelo, CR_t, Classificação da Rentabilidade do Empresa em t, conforme a divisão detalhada na Tabela 1.

Tabela 1. Definição da classificação do valor da rentabilidade do patrimônio ajustado

CR_t	Condição	Avaliação
1	$VR_t < 2{,}85\%$	Retorno de rentabilidade ruim
2	$2{,}85\% \leq VR_t < 12{,}5\%$	Retorno de rentabilidade médio
3	$VR_t \geq 12{,}5\%$	Retorno de rentabilidade ótimo

As classes foram definidas de modo que todas tivessem, aproximadamente, o mesmo número de observações (por volta de 740).

- Variáveis independentes: 11 parâmetros, em $t - 1$, ou seja, dos anos de 1996 a 2005.
- Observações: cada observação foi composta com a variável dependente CR_t, Classificação da Rentabilidade da Empresa, no ano atual (em t) e 11 parâmetros da mesma empresa (variáveis independentes) no ano anterior (em $t - 1$).

O conjunto total de observações foi separado em dois subconjuntos:

- Subconjunto de treinamento: 2.215 observações com dados de empresas entre 1996 e 2004 com classificação de rentabilidade de 1997 a 2005.
- Subconjunto de verificação: 410 observações com dados de empresas em 2005 com classificação de rentabilidade em 2006.

2.2. Definição do Modelo de Regressão Linear Múltipla (RLM)

Para selecionar as variáveis de entrada do modelo RLM, utilizamos o Método de Eliminação, utilizando coeficiente de significância $\alpha = 5\%$. Para comparar os modelos, foram utilizadas as estatísticas R^2 Ajustado e F. A Tabela 2 exibe o conjunto escolhido no final do processo.

Tabela 2. Coeficientes e estatísticas do modelo RLM escolhido

CONJUNTO FINAL	Coeficientes	Estatística-T	P-Valor
Constante	1,5521	31,25269	0,00000
CCL_{t-1} – Capital Circulante Líquido	–0,00014	–2,30015	0,02153
LLA_{t-1} – Lucro Líquido Ajustado	0,00026	3,91740	0,00009
PLA_{t-1} – Patrimônio Líquido Ajustado	–0,00004	–3,40715	0,00067
EG_{t-1} – Endividamento Geral	0,00536	6,51043	0,00000
MV_{t-1} – Margem de Vendas	0,02229	15,18037	0,00000
IGA_{t-1} – Índice de Giro de Ativo	0,07252	5,29801	0,00000
VR_{t-1} – Valor da Rentabilidade	0,00186	4,77646	0,00000

O RLM é um modelo contínuo. Desta forma, para obter a classificação dos rendimentos, efetuamos o arredondamento (para o inteiro mais próximo) dos valores de saída calculados pelo modelo. Para este conjunto, a taxa de acerto do modelo RLM foi de 37,65% (834 acertos em 2215 observações). A taxa de acerto do modelo RLM para os dados de treinamento (37,65%) ficou muito próxima à taxa de acerto teórica de um sorteio aleatório (33%), dada a composição do conjunto de dados de treinamento

com, aproximadamente, um terço das observações para cada classe. Este baixo resultado também foi indicado pelo baixo valor da estatística R^2 Ajustado (0,1683).

Apesar do baixo percentual de acerto, podemos utilizar os coeficientes calculados pelo modelo para interpretar o fenômeno sob a condição 'Ceteris Paribus'. O modelo estimou coeficientes positivos para cinco variáveis: LLA_{t-1}, EG_{t-1} , MV_{t-1}, IGA_{t-1} e VR_{t-1}. Isto indica que aumentos nos valores dessas variáveis contribuiriam para melhorar a classificação das empresas. Nesse conjunto, a variável que mais contribuiu para a melhora da classificação das empresas foi a IGA_{t-1}, índice que mede eficiência operacional, o que estaria de acordo com a intuição econômica de que, uma empresa mais eficiente é mais rentável. As contribuições positivas das demais variáveis são suportadas pela teoria de finanças corporativas e contabilidade empresarial, na fórmula de *ROE – Return on Equity* apresentada em Brealy, Myers & Allen [19] e Marion [20]. A variável PLA_{t-1} aparece no denominador de fórmulas de rentabilidade o que justifica o coeficiente negativo estimado pelo modelo para essa variável. No caso da variável CCL_{t-1}, que representa o total de recursos de curto prazo disponíveis para financiamento das atividades da empresa, um aumento dessa variável poderia indicar melhora na capacidade de liquidação de compromissos de curto prazo, menor risco de insolvência, e portanto, favorecer a rentabilidade da empresa. Porém, de acordo com Brealy, Myers & Allen [19], a relação entre capital circulante líquido e rentabilidade também depende da composição do ativo circulante da empresa, que é formado por: estoque e contas a receber. Sendo assim, a teoria indica que a influência do capital circulante líquido sobre a rentabilidade depende de características da gestão de capital de giro das empresas. Para este experimento, o modelo de RLM estimou um coeficiente negativo para a variável CCL_{t-1}. Em valor absoluto, é o coeficiente de menor valor, portanto o menos influente, o que reforça a dualidade descrita pela teoria.

2.3. Definição do Modelo de Regressão Logística Ordenada (RLO)

Para o modelo RLO, também foi utilizado o Método de Eliminação, com coeficiente de confiabilidade α = 5%. Para comparar os modelos, foi utilizado o critério de informação de Schwarz. A Tabela 3 exibe o conjunto escolhido no final do processo.

Tabela 3. Coeficientes e estatísticas do modelo RLO escolhido

CONJUNTO FINAL	Coeficientes	Estatística–T	P–Valor
$\alpha 1$ valor de corte 1	0,60057	4,25168	0,00002
$\alpha 2$ valor de corte 2	2,27927	15,29821	0,00000
LLA_{t-1} – Lucro Líquido Ajustado	–0,00228	–3,14480	0,00166
PLA_{t-1} – Patrimônio Líquido Ajustado	0,00010	3,50202	0,00046
EG_{t-1} – Endividamento Geral	–0,01523	–6,69247	0,00000
MV_{t-1} – Margem de Vendas	–0,06592	–8,98779	0,00000
IGA_{t-1} – Índice de Giro de Ativo	–0,17925	–4,72320	0,00000
VR_{t-1} – Valor da Rentabilidade	–0,02088	–5,71044	0,00000

O modelo RLO pertence à classe de modelos de regressão capazes de tratar problemas com variáveis dependentes discretas. Para este conjunto, a taxa de acerto do modelo RML foi de 54,54% (i.e. 1208 acertos em 2215 observações). Este resultado foi obtido com um conjunto menor de variáveis independentes (todas as variáveis selecionadas pelo RLM menos a CCL_{t-1}). Isto reforça a indicação de que a capacidade de generalização do modelo RLO é superior à do RLM. O modelo RLO estimou coeficiente positivo apenas para a variável *PLA*. Neste tipo de modelo, o sinal positivo indica que o aumento no valor dessa variável contribui para piorar a classificação das empresas, e o sinal negativo indica que o aumento no valor dessa variável contribui para melhorar a classificação das empresas. Ou seja, o modelo RLO apontou as mesmas relações encontradas pelo modelo RLM, com exceção da variável CCL_{t-1}. O resultado produzido indica que o modelo RLO não encontrou evidências para considerar essa variável relevante, pois, conforme o exposto no item anterior, a rentabilidade pode ser influenciada pelo capital circulante líquido, tanto positiva quanto negativamente.

2.4. Definição do Modelo de Redes Neurais (RN)

A teoria de RN não provê critérios para seleção de variáveis de entrada, semelhantes aos utilizados nos itens anteriores. Desta forma, a arquitetura de uma RN acaba sendo escolhida com base na experiência e conhecimento dos modeladores. Essa experiência acumulada aparece na literatura especializada [17], [21] na forma de um conjunto de recomendações que podem ser utilizadas para orientar o processo de definição da arquitetura de uma RN, quais sejam:

- Normalizar as variáveis de entrada e saída de forma que todas fiquem em um mesmo intervalo específico [–1, 1];
- Testar arquiteturas com diferentes números de neurônios na camada escondida;
- Executar vários treinamentos com pontos iniciais diferentes.

Para que a RN pudesse devolver a classificação da empresa e não um número real, definimos a RN com *output layer* com três neurônios: um para cada classe, representando seu estado de ativação, de acordo com o quadro abaixo onde 1 indica o mais alto estado de ativação e –1, o mais baixo. Para a construção das RN, a variável dependente foi recodificada para essa forma, conforme detalhado na tabela a seguir (Tabela 4).

Tabela 4. Codificação para o Output Layer do modelo de RN.

Classe 1	Classe 2	Classe 3	Classificação Observada (CR_t)	Avaliação
1	–1	–1	1	Retorno de rentabilidade ruim
–1	1	–1	2	Retorno de rentabilidade médio
–1	–1	1	3	Retorno de rentabilidade ótimo

No caso do experimento desenvolvido neste estudo, para as características gerais da arquitetura da RN, adotamos configurações consideradas convencionais [22]:

- Tipo da Rede: acíclica;
- Número de camadas: três (de entrada, escondida e de saída);
- Camada de entrada: com número de neurônios igual ao número de variáveis de entrada, com pré–processamento de normalização de dados;
- Camada de saída: com três neurônios (número de saídas);
- Função de transferência: *tansig*, para as camadas de entrada e escondida, e *purelin*, para a camada de saída.

Para os parâmetros de treinamento, os valores foram arbitrados:

- Valor objetivo: 1000;
- Número máximo de execuções: 100;
- Tamanho dos subconjuntos:
  - Treinamento (usado para calibrar a RN): 50%;
  - Validação (usado para evitar o *overfittting*): 20%;
  - Teste (usado para comparar modelos): 30%.

Com relação à função de treinamento, para evitar a necessidade de arbitrar o número de neurônios na camada escondida ou ter que testar arquiteturas com diferentes números de neurônios, escolhemos a função de treinamento baseada no método de Regularização Bayesiana que, de acordo com o trabalho desenvolvido por MacKay [23], e detalhado por Foresee e Hagan [24], estima os parâmetros ótimos de regularização de uma RN. Reguralização é o nome dado ao processo de otimização que tem como objetivo encontrar valores de pesos e *bias* que minimizam a função de desempenho. A Regularização Bayesiana, com o objetivo de evitar o *overfitting* e melhorar a capacidade de generalização da RN, acrescenta ao processo de otimização um componente que penaliza a função com o aumento do número de coeficientes, o que permite que o número de neurônios da camada escondida seja redefinido automaticamente durante o processo de treinamento.

Trabalhar todos os conjuntos possíveis gerados pelas 11 variáveis de entrada significaria testar um grande número de combinações, mais exatamente 2047, todas as permutações possíveis com 1, 2 até 11 elementos. No entanto, aproveitou-se do fato de que os modelos RLM e RLO já haviam apontado as variáveis com relações lineares significativas para tomá–las como base de composição do conjunto de variáveis de entrada e, desta forma, ao invés de 2047, apenas 17 conjuntos foram considerados:

- Conjunto selecionado pelo modelo RLO;
- Conjunto selecionado pelo modelo RLM;
- 15 conjuntos gerados pela combinação das outras 4 variáveis acrescidas ao conjunto selecionado pelo modelo RLM.

Para determinar o conjunto de variáveis de entrada e a arquitetura da RN, foram realizadas 100 execuções para cada uma das 17 combinações. Para fins de comparação, utilizou-se do número de acertos no subconjunto de teste. A lista a seguir exibe o conjunto escolhido no final do processo.

- CCL_{t-1} – Capital Circulante Líquido em t – 1;
- LLA_{t-1} – Lucro Líquido Ajustado em t – 1;
- PLA_{t-1} – Patrimônio Líquido Ajustado em t – 1;
- RCE_{t-1} – Riqueza Criada por Empregado em t – 1;
- CV_{t-1} – Crescimento em Vendas em t – 1;
- EG_{t-1} – Endividamento Geral em t – 1;
- MV_{t-1} – Margem de Vendas em t – 1;
- IGA_{t-1} – Índice de Giro de Ativo em t – 1;
- VR_{t-1} – Valor da Rentabilidade em t – 1.

Para cada observação, o modelo de RN produziu três valores correspondentes ao estado de ativação de cada classe. Para obter a classificação da rentabilidade das empresas, analogamente à construção exibida na Tabela 4, escolheu-se a classe com valor de ativação mais alto, conforme o exemplo na Tabela 5 a seguir.

Tabela 5. Tratamento do resultado do modelo de RN

Classe 1	Classe 2	Classe 3	Classificação Observada (CR_t)	Avaliação
0,880	–0,830	–1,050	1	Retorno de rentabilidade ruim
–0,990	0,704	–0,715	2	Retorno de rentabilidade médio
–0,774	–0,731	0,505	3	Retorno de rentabilidade ótimo

Para este conjunto, a taxa de acerto do modelo de RN foi de 60,59% (1342 acertos em 2215 observações). O resultado do experimento indicou que o modelo de RN conseguiu captar mais relações entre as variáveis que os modelos RLM e RLO, com a inclusão das variáveis RCE_{t-1} e CV_{t-1} e, desta forma, melhorar o percentual de acertos. A Tabela 6 mostra que o modelo de RN produziu o melhor resultado para os dados de treinamento.

Tabela 6. Comparação dos resultados dos modelos para dados de treinamento.

Dados de Treinamento	Modelo		
	RLM	RLO	RN
Acertos	834	1.208	1.329
% Acertos	37,65%	54,54%	60,00%

3. Resultados

Para comparar a capacidade de generalização dos modelos, os dados de verificação foram utilizados para estimar a classe de rentabilidade de 410 empresas em 2006, a partir de informações de 2005, e obtendo-se os resultados exibidos na Tabela 7.

Tabela 7. Comparação entre os resultados dos modelos para dados de verificação.

Dados de Verificação	RLM	RLO	RN
Acertos	144	245	290
% Acertos	35,12%	59,76%	70,73%

Para os dados de verificação, a taxa de acerto do modelo RLM de 35,12% (144 acertos em 410 observações) ficou abaixo da taxa obtida com os dados de treinamento (37,65%). Esta baixa capacidade de generalização indica que esse tipo de modelo não se mostrou adequado para tratar o tipo de problema apresentado neste experimento. O resultado do modelo de RN para os dados de verificação foi expressivamente superior ao do RLO pois, além de partir de uma taxa de acertos mais alta, o aumento do número de acertos foi quase duas vezes maior que o observado no RLO. Isto é, a taxa de acertos da RN passou de 60,59% para 70,73%, subindo 10,14% (contra os 5,22% do modelo RLO). Este resultado fornece indícios de que o modelo RN, quando comparado aos modelos RLM e RLO, conseguiu captar melhor o relacionamento entre as variáveis de entrada e saída, sem perder a capacidade de generalização, conforme o resumo dos resultados comparativos na Tabela 8.

Tabela 8. Comparação percentual entre os resultados dos modelos RLM e RLO com o RN para dados de verificação.

Modelo Base de Comparação	Aumento % Absoluto do modelo RN	Aumento % Relativo do modelo RN
RLM	35,61%	101,39%
RLO	10,98%	18,37%

4. Discussão

De acordo com a teoria de RN [22], não existem procedimentos para a extração das regras utilizadas pela RN para relacionar efeitos de variações nas variáveis de entrada sobre as variáveis de saída. Isto poderia levantar a suspeita de que o número de acertos mais alto resultante do modelo de RN tenha sido apenas uma coincidência.

Para verificar esta questão, examinou-se a distribuição dos percentuais de acerto para os dados de treinamento nas 1700 execuções realizadas (100 execuções para cada uma das 17 combinações) e contou-se o número de resultados da RN que ficaram abaixo do percentual de acertos do modelo RLO (54,54%). A Figura 1 mostra os seguintes números: 7 resultados do modelo de RN ficaram abaixo do modelo RLO (aprox. 0,5%) e 1693 acima (aprox. 99,5%).

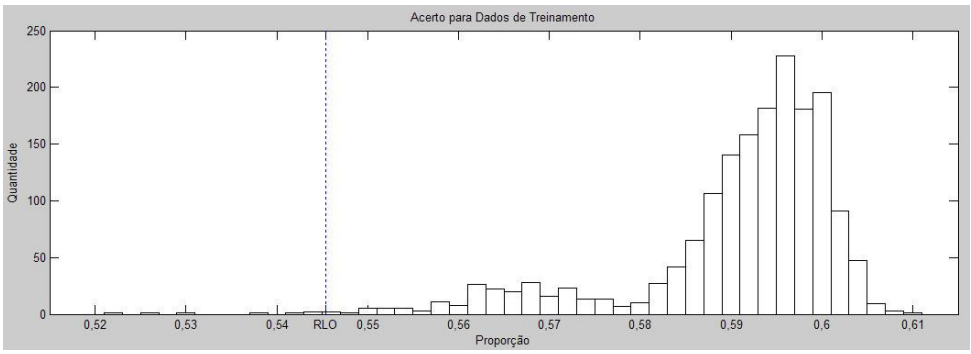

Figura 1 – Distribuição dos percentuais de acerto para os dados de treinamento.

Da mesma forma, com a finalidade de verificar a consistência do resultado obtido, o mesmo procedimento foi executado para os dados de verificação nas 1700 execuções realizadas. A Figura 2

mostra a contagem de quantos resultados da RN ficaram abaixo do resultado dado pelo modelo RLO (59,76%): 255 resultados do modelo RN inferiores ao do modelo RLO (3,75%) e 6545 superiores (96,25%).

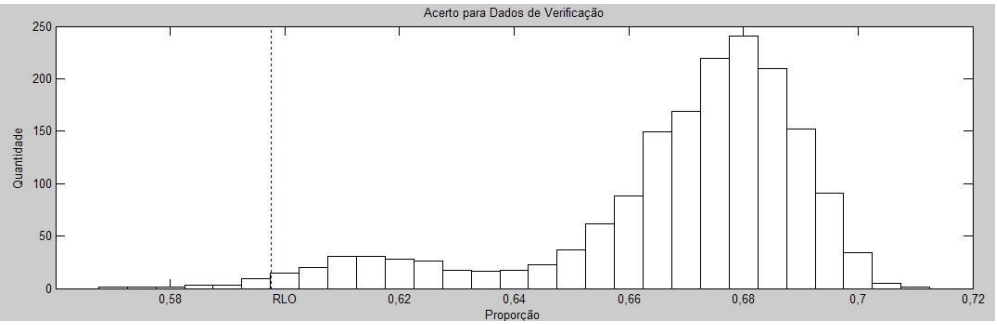

Figura 2 – Distribuição dos percentuais de acerto para os dados de verificação.

Verificou-se desta forma que, mesmo considerando qualquer conjunto de variáveis de entrada, os modelos de RN foram, em sua maioria, capazes de gerar resultados superiores aos do modelo RLO.

5. Conclusões

Os experimentos descritos neste trabalho fornecem evidências de que, para a estimação de classificação de rentabilidade de empresas, modelos de redes neurais podem gerar resultados melhores do que modelos mais comumente empregados como os de regressão linear e de logística ordenada. Este estudo ilustra uma aplicação prática de técnicas de classificação de padrões que podem ser utilizadas para desenvolver e aprimorar qualquer tipo de ferramenta de apoio à tomada de decisões, em particular relativos à escolha de investimentos, bem como, fazer parte de sistemas de gerenciamento de risco, avaliação de crédito e identificação de comportamentos.

Para melhorar os percentuais de acertos em relação ao conjunto de informações analisadas, estudos semelhantes podem ser realizados levando–se em conta empresas por setor de atividade. Outra possibilidade seria acrescentar variáveis que, de alguma forma pudessem expressar influências de fatores macroeconômicos, ou mesmo que conseguissem quantificar subjetividade.

Sobre a escolha das variáveis de entrada do modelo de redes neurais, seria interessante explorar o uso de procedimentos estatísticos, como por exemplo, a medida de informação mútua.

Quanto às redes neurais em si, outros tipos, como por exemplo, *Localist Attractor Network*, que trabalha com variável dependente discreta, ou *Self–Organization Map*, que utiliza método de aprendizado não supervisionado poderão ser testados.

Adicionalmente, destacamos que os resultados dos modelos RLM e RLO não serviram apenas como marcas comparativas, mas também foram fundamentais para facilitar a definição do conjunto de variáveis de entrada dos modelos de RN. Desta forma, os experimentos descritos também demonstram que o uso combinado de técnicas pode ajudar a produzir melhores resultados.

Referências:

[1] D. Baestaens, W. Bergh, D. Wood, Neural network solutions for trading in financial markets – **London Pitman** (1994).

[2] Y. Abu–Mostafa, A. Atiya, Introduction to financial forecasting – Applied Intelligence, 6(3):205–213, **Kluwer Academic Publishers** (1996).

[3] E. Gately, Neural Networks for Financial Forecasting – **John Wiley & Sons** (1995).

[4] E.K. Jason, Forecasting Financial Markets Using Neural Networks: an analysis of methods and accuracy – **Naval Postgraduate School, Monterey**, Califórnia/EUA (1998).

[5] M. Medeiros, A. Veiga, C.E. Pedreira, Modelling Exchange Rates: Smooth Transitions, Neural Networks, and Linear Models – **IEEE Transactions on Neural Networks,** 12, 755–646 (2001).

[6] A. Chen, M. Leung, – Performance Evaluation of Neural Network Architectures: The Case of Predicting Foreign Exchange Correlations – **Journal of Forecasting**, Vol. 24 (2006).

[7] P. Sahoo, P. Hathy, Prediction of Mutual Funds: Use of Neural Network Technique – **The Icfai Journal of Applied Finance**, Vol 13–11 (2007).

[8] K. Chung, S. Tan, D. Holdsworth, Insolvency Prediction Model Using Multivariate Discriminant Analysis and Artificial Neural Network for the Finance Industry in New Zealand – **International Journal of Business and Management,** Vol. 39–1 (2008).
[9] E. Lemos, M. Steiner, J. Nievola, Análise de crédito bancário por meio de redes neurais e árvores de decisão : uma aplicação simples de data mining – **Revista de Administração,** São Paulo – Vol. 40–3 (2005).
[10] S. Freitas, A. Souza, Utilização de redes neurais na precificação de opções – **Resenha BM&F**, Vol.150 (2002).
[11] G. Chagas, Precificação de opções de dólar no mercado brasileiro utilizando redes neurais e algoritmos genéticos – **Escola de Economia de São Paulo**, Fundação Getúlio Vargas (2006).
[12] C. Pinheiro, A. Evsukoff, N Ebecken, Revenue Recovering with Insolvency Prevention on a Brazilian – **ACM SIGKDD Explorations Newsletter**, Vol. 8–1(2006).
[13] A. Arnauld, R.Cunha, T. Ludermir, P. Adeodato, Modelo Híbrido com Redes Neurais Artificiais e Técnicas não–supervisionadas para o problema de credit scoring – **XXV Congresso da Sociedade Brasileira de Computação** (2005).
[14] L. Zaneti, F. Almeida, Exploração do uso de redes neurais na previsão do comportamento de ativos financeiros – **III SEMEAD, FEA/USP** (1998).
[15] F. Almeida, P. Dumontier, O uso de redes neurais em avaliação de riscos de inadimplência – **Revista de Administração FEA/USP**, Vol.31–1 (1996).
[16] M. Lam, Neural Network techinques for financial performance prediction: integrating fundamental and techincal analysis – **Decision Support Systems,** Vol 37–4 (2004).
[17] R. Duda, D. Stork, P. Hart, Pattern Classification – **John Wiley & Sons** (2005).
[18] Revista Exame, Melhores e Maiores – Editora Abril (2007).
[19] R. Brealy, S. Myers, F. Allen, Principles of Corporate Finance, 8th Edition – **Irwin/McGraw–Hill** (2006).
[20] J.C. Marion, Análise das Demonstrações Contábeis: contabilidade empresarial – 2a Edição – **Editora Átlas**, São Paulo (2002).
[21] C.Bailer–Jones, Introduction to Machine Learning and Pattern Recognition – **Lecture Hörsaal, Max–Planck, Institut für Astronomie** Fev/Mar (2008).
[22] M. Hagan, H. Demuth, M. Beale, Neural Network Design – **PWS Publishing** (1996).
[23] D.J.C. MacKay, Bayesian Interpolation – **Neural Computation**, Vol. 4–3 – pp. 415–447 (1992).
[24] F.D. Foresee, M.T. Hagan, Gauss–Newton approximation to Bayesian regularization – **Proceedings of the 1997 International Joint Conference on Neural Networks**, pp. 1930–1935 (1997).