

Empresas Criadoras x Destruidoras de Valor: um estudo do setor de Siderurgia e Metalurgia

RESUMO

Em qualquer economia, o preceito objetivo de criação de valor é construção de empresas competitivas, promovendo oportunidades melhores de crescimento econômico. O objetivo desse estudo é avaliar quais as variáveis que discriminam as empresas criadoras de valor das empresas destruidoras de valor, baseado em dados empíricos. Para atingir o objetivo do trabalho foram coletadas as demonstrações contábeis de 38 empresas do setor de metalurgia e siderurgia para o período de 1996-2003 e foram determinadas as principais medidas econômicas de criação de valor ajustadas ao Brasil. Os dados foram analisados por Análise Discriminante e Redes Neurais Artificiais. Os resultados mostram a dificuldade de trabalhar com dados não lineares em Análise Discriminante e algumas vantagens da Rede Neural na análise dos dados. A Análise Discriminante gerou um resultado incapaz de classificar corretamente as empresas e a Rede Neural não fornece as variáveis adotadas pelo modelo gerado, apesar de classificar relativamente bem as empresas. Ressalta-se a importância de estudos que aumentem a capacidade preditiva dos investidores quanto à identificação de empresas criadoras de valor.

1. Introdução

O conceito de valor econômico agregado foi introduzido com maior ênfase, pelas companhias em todo mundo, principalmente a partir de 1980. Diante da globalização e abertura dos mercados, o conceito de uma gestão baseada no valor foi adotado visando a reduzir as ineficiências das empresas, tornando-as mais competitivas e valorizadas no mercado. Nesse processo de internacionalização dos negócios e abertura econômica, ficou amplamente entendido o objetivo da teoria de Finanças de maximização da riqueza dos investidores (acionistas).

Indicadores tradicionais baseados nos lucros contábeis foram substituídos, nesse modelo de gestão, por medidas indicativas do sucesso econômico que incorporam o valor agregado aos acionistas. A análise econômica do valor revela a dinâmica de um negócio de forma menos sintética que qualquer outro instrumento de análise.

As principais contribuições do modelo de avaliação baseado em medidas econômicas são resumidas a seguir: (i) introduz uma taxa de risco para confrontar com a rentabilidade apurada e descontar os fluxos de caixa gerados. O modelo admite que os investidores fixem sua atenção nos resultados operacionais de caixa esperados de serem realizados no futuro e no risco envolvido; (ii) o modelo aproxima bastante a gestão das companhias com a racionalidade financeira de avaliação de seus proprietários. O foco no mercado permite que os executivos criem riqueza para si mesmos e também para os acionistas; (iii) é possível na avaliação identificar-se os fatores que agregam ou destroem valor econômico, permitindo uma compreensão crítica das estratégias adotadas nos negócios; (iv) o modelo econômico demonstra a ineficiência de se avaliar o desempenho da empresa centrado em medidas contábeis tradicionais, como o lucro por ação.

Apesar de a medida de valor econômico agregado ser muitas vezes considerada como moderna, o conceito é bastante antigo, sendo conhecido na literatura contábil e financeira por lucro residual ou lucro econômico. Uma referência do conceito encontra-se em trabalho do economista Alfred Marshall (1890, v. 1, p. 142)¹, ao expressar o efetivo

significado de lucro dos acionistas (lucro líquido) como o valor que resta do lucro após deduzir-se o custo corrente de seu capital (capital próprio).

O conceito básico inserido no lucro econômico é que uma empresa somente apura um resultado positivo se suas receitas de vendas cobrirem todas as suas despesas, inclusive a expectativa mínima de remuneração do capital investido pelos acionistas.

A estratégia de lucro econômico inclui alguns fatores, ou direcionadores de valor, fundamentais para o processo de criação de valor. Podem ser citados, entre outros, a estrutura e custo de capital, determinantes do risco operacional e financeiro da empresa, retorno sobre o investimento (ROI) e oportunidades de crescimento que adicionam valor.

As empresas vêm direcionando sua atenção nas últimas décadas, de forma privilegiada, na maximização dos fluxos operacionais de caixa e retorno de seus investimentos pela utilização de diversas estratégias financeiras e corporativas. Os negócios passam, então, a serem avaliados pelas métricas de valor econômico agregado, mantendo uma estreita sintonia com as expectativas dos investidores de mercado de maximizar a riqueza do capital investido.

As medidas de criação de riqueza são atualmente demandadas pelo mercado, e consagradas na teoria de Finanças, como indicadores fundamentais de avaliação de desempenho e criação de riqueza. O foco do valor econômico para o acionista desafia as empresas a produzirem um *spread* econômico cada vez maior, medido pela diferença entre a taxa de retorno e o custo de oportunidade do capital investido.

Assim, **o objetivo do estudo é avaliar quais as variáveis que discriminam as empresas criadoras de valor das empresas destruidoras de valor.**

2. Criação de valor econômico nas empresas

A maximização do preço de mercado das ações predominou inicialmente nas decisões corporativas dos Estados Unidos, espalhando-se posteriormente para a Europa, América Latina e Ásia. As razões que levaram as várias corporações mundiais a priorizar a maximização da riqueza como seu principal objetivo foram as mesmas descritas. A aceitação dessa abordagem nas demais partes do mundo, no entanto, foi mais lenta que nos Estados Unidos, explicada principalmente pelo protecionismo desses países em relação às suas grandes empresas nacionais.

Essa barreira protecionista começou a recuar somente com o início do processo de globalização, acompanhada de uma crescente integração econômica verificada nos diversos países. Os mercados comerciais e financeiros foram liberalizados, mais acentuadamente a partir da década de 80, e as empresas passaram a conviver com um maior grau de concorrência.

Nesse ambiente de maior competitividade, as corporações desses países tiveram sua viabilidade econômica e continuidade mensurada também pela criação de valor. Desempenhos fracos, em termos de remuneração do capital investido, determinam mudanças rápidas dos recursos para corporações agregadoras de valor.

A resposta das empresas fora dos EUA a essas fortes mudanças de mercado veio, de forma idêntica, por meio de métricas de desempenho baseadas no valor agregado. Foi uma forma de satisfação dessas empresas à crescente demanda do mercado por oportunidades de investimentos com potencial de maximização de riqueza. Os administradores passaram a priorizar desempenho, pois é assim que são cobrados e avaliados pelos seus acionistas.

O número de empresas no mundo que vêm se adaptando a essas novas exigências de elevar o valor econômico para o acionista é crescente. Essa preocupação é mais evidente

em companhias listadas em bolsas de valores de expressão internacional (NYSE – New York Stock Exchange, por exemplo). Em verdade, essas corporações entenderam que economias voltadas aos acionistas são capazes de produzir maior riqueza em relação a outros modelos que não privilegiam o valor para o acionista nas várias decisões corporativas.

O fundamento básico do objetivo de criação de valor, em qualquer economia, é a construção de empresas mais fortes e competitivas, capazes de promover padrões de vida mais elevados e melhores oportunidades de crescimento econômico para as nações. Os acionistas são os verdadeiros proprietários das empresas, e o objetivo de toda administração é o de maximizar o seu valor de mercado.

3. Metodologia e Análise dos Resultados

Para atingir o objetivo do trabalho foi usada pesquisa documental para coleta de dados e a análise dos dados foi feita com o auxílio de ferramentas estatísticas. Foram coletadas as demonstrações contábeis de 38 empresas do setor de metalurgia e siderurgia conforme disponibilizadas pela Economáticaⁱⁱ e foram determinadas, a partir desses conjuntos de informações contábeis, as principais medidas econômicas de criação de valor ajustadas ao Brasil. As medidas foram calculadas em cada empresa para cada um dos oito anos considerados na pesquisa (1996-2003).

As ferramentas adotadas para análise dos dados foram: Análise Discriminante e Redes Neurais. Para que tais ferramentas pudessem discriminar entre as variáveis, foi necessário definir um grupo de variáveis para inserir no banco de dados. As variáveis foram selecionadas com base nos direcionadores de valor propostos na bibliografia pesquisada (Stern e Shiely, 2001; Rappaport, 2001; Copeland *et al.*, 2000; Martin e Petty, 2004; Assaf, 2003a; Araújo, 2002; Bonízio, 2005) e estão dispostas no quadro 1.

Variável	Nome nas tabelas	Descrição
Lucro Líquido	lucroliq	valor obtido das Demonstrações publicadas pelas empresas
Lucro Operacional Ajustado	LOpAjust	apuração do resultado líquido do Imposto de Renda, considerando receitas financeiras, outras receitas e despesas operacionais e resultado não operacional; o resultado operacional apurado encontra-se demonstrado em moeda de final de exercício.
Endividamento	pplmd	Passivo Oneroso dividido pelo Patrimônio Líquido
Beta alavancado	Betaalavmd	risco sistemático da empresa ou setor (coeficiente)
Custo de Capital Próprio	ke	remuneração mínima esperada pelos proprietários de capital em relação à entidade investida.
Custo médio Ponderado de Capital	wacc	custo médio ponderado do capital é a rentabilidade esperada de uma carteira constituída por todos os títulos da empresa.
Retorno sobre o PL	roe	o retorno sobre o patrimônio líquido ou <i>return on equity</i> (ROE), relaciona o ganho obtido pelos sócios e o investimento realizado na empresa.
Retorno sobre o Investimento	roi	é o resultado da relação entre o resultado líquido do período e o total de investimentos. O <i>return on investment</i> (ROI) em sua versão mais simplificada pode ser obtido pela divisão do resultado líquido pelo ativo total.
Margem Líquida	margem	é o valor do Lucro após o Imposto de Renda dividido pela Receita de Vendas Líquida ou, diferença entre o preço do produto e todos os custos e despesas envolvidos na fabricação.

Giro do ativo	giro	é a receita bruta de vendas dividida pelo ativo total ajustado; mede a eficiência operacional da empresa e deve ser comparado com a margem de lucro sobre vendas.
Taxa de Reinvestimento	TxReinvest.b	A taxa de reinvestimento dos resultados operacionais, medida relevante na avaliação, é mensurada pela relação entre os reinvestimentos líquidos e o lucro operacional após o imposto de renda, ou seja: Reinvestimentos Líquidos = (Investimentos em Capital Fixo + Variações no Giro) – Depreciação ; Taxa de Reinvestimento (b) = Reinvestimentos Líquidos / Lucro Operacional após IR.
EVA ^{®iii}	EVA [®]	valor econômico agregado; EVA [®] = LOP – (CMPC– INVT)

Quadro 1 – Grupo de Variáveis utilizadas na pesquisa

Fonte: elaborado pelos autores com base em Stern e Shiely (2001), Rappaport (2001), Copeland *et al.* (2000), Martin e Petty (2004), Assaf (2003a), Araújo (2002) e Bonízio (2005)

A coluna ‘Nome nas Tabelas’ do quadro 1 indica como a variável aparece nas tabelas de resposta do *software* estatístico. A seguir serão analisadas as duas ferramentas utilizadas para seleção das variáveis que melhor discriminam as empresas criadoras das destruidoras de valor.

3.1 Análise Discriminante

A análise discriminante é uma técnica estatística capaz de identificar as características (variáveis) que melhor diferenciam os grupos de observações. Ela pode ser definida como uma combinação linear de duas (ou mais) variáveis independentes que discriminarão melhor entre grupos definidos *a priori* (HAIR *et al.*, 2005). Para realizar os cálculos e os testes estatísticos foi utilizado o *software SPSS*.

A função discriminante pode ser descrita assim:

$$Z_{jk} = a + w_1 \cdot X_{1k} + w_2 \cdot X_{2k} + \dots + w_n \cdot X_{nk}$$

onde:

Z_{jk} = escore discriminante

a = coeficiente linear

w_i = peso discriminante (ponderação)

x_{ik} = coeficiente discriminante (variável independente)

Como a análise discriminante é uma ferramenta apropriada nos casos em que a variável dependente é categórica e as variáveis independentes são métricas, foi inserida no grupo de variáveis uma variável categórica (chamada “z”) que classifica a empresa como criadora (1) ou destruidora de valor (0). As observações (foram 304 observações: 38 empresas em 8 anos) com EVA[®] positivo foram classificadas como criadoras de valor e as observações com EVA[®] negativo, como destruidoras de valor.

Para que o programa realize a análise discriminante é preciso selecionar a variável dependente (categórica) e as independentes (métricas). Assim, a variável dependente selecionada foi “z” e as variáveis expostas no quadro 1 foram colocadas como variáveis independentes (com exceção do EVA[®] por esta ser a própria definição da variável dependente e não um direcionador de valor). O método de análise selecionado foi o método *Stepwise* que inclui e exclui variáveis do modelo automaticamente conforme sua capacidade de aumentar o poder explicativo do modelo.

No método “Stepwise” a análise começa sem nenhuma variável e nos passos seguintes as variáveis são adicionadas ou removidas consoante a sua contribuição para a capacidade discriminante da 1ª função discriminante. Assim, uma variável

presente na função discriminante é removida se a capacidade da função discriminante não for significativamente reduzida pela remoção dessa variável, e no outro extremo, uma variável é adicionada se aumentar significativamente a capacidade discriminativa da função discriminante (MAROCO, 2003, p. 359)

O critério de inclusão/remoção das variáveis foi a Distância de Mahalanobis com significância no teste F de 5% para entrada de variáveis e de 10% para saída. De acordo com Hair *et al.* (2005 p. 222), se o método utilizado for o *Stepwise*, este critério é um dos mais adequados. A Distância de Mahalanobis entre as variáveis i e j é dada por (MAROCO, 2003, p. 360)

$$DM_{ij} = \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)} .$$

Contudo, antes de executar a análise discriminante ressalta-se ser “recomendável atender a certas condições para uma aplicação apropriada da análise discriminante” (HAIR *et al.*, 2005, p. 220). Estas condições são: normalidade multivariada das variáveis independentes; linearidade das relações; falta de multicolinearidade entre as variáveis independentes; ausência de *outliers*; e matrizes de variância-covariância homogêneas.

Em um primeiro momento, os dados foram analisados com o grupo completo de observações: 304 casos em 12 variáveis. Entretanto os resultados muitas vezes não foram calculados devido a grande quantidade de dados perdidos (*missing values*). Assim, optou-se por eliminar todos os casos em que o EVA[®], variável de maior interesse por caracterizar a variável dependente z , fosse um dado perdido. Assim, foram eliminados 74 casos, o que representa aproximadamente 25% da amostra inicial. Esta quantidade de dados perdidos é explicada por um grande número de empresas que não apresentava dados em todos os anos analisados, seja por não possuir capital aberto nos primeiros anos, ou por ter sido fechada ou incorporada nos últimos anos, ou por motivos diversos. Portanto, a amostra efetivamente analisada neste tópico era composta por 230 casos e doze variáveis independentes, além do EVA[®].

Assim, os dados foram analisados para verificar se essas condições eram atendidas. Foram utilizados também os softwares *Eviews* e *BestFit* para complementar e/ou verificar tais análises. Em resumo, os resultados dessa pré-análise foram:

1. nenhuma das variáveis independentes apresenta distribuição normal – isto pode ser observado nas tabelas 1 e 2 que apresentam o Teste Kolmogorov-Smirnov (aderência dos dados à curva normal) com todos os p -values (‘Sig’ nas tabelas) com valor zero, ou seja, para os níveis de significância tradicionais (1%, 5% e 10%) a hipótese de normalidade deve ser rejeitada;
2. não existe linearidade em todas as relações e há muitos *outliers* em diversas variáveis – isto pode ser observado na figura 1 que traz os gráficos de todos os pares de variáveis;
3. há falta de multicolinearidade entre as variáveis, como pode ser observado na tabela 3, pois existem poucas correlações superiores a +0,5 ou inferiores a -0,5 (a alta correlação se aproxima de 1 ou -1).

Contudo, ainda falta analisar um dos principais pressupostos da análise discriminante: a homogeneidade das matrizes de variância-covariância. Este pressuposto será testado quando a análise discriminante for gerada no SPSS. Portanto, por enquanto, apenas o pressuposto da falta de multicolinearidade foi atendido.

One-Sample Kolmogorov-Smirnov Test							
		lucroliq	LOpAjust	eva	pplmd	Betaalavmd	ke
N		230	230	230	228	223	230
Normal Parameters(a,b)	Mean	98824,6	120994,5	-52684,6	2,6	1,6	,27
	Std. Deviation	309871,7	285378,4	195520,4	21,5	8,2	1,2
Most Extreme Differences	Absolute	,301	,288	,244	,437	,437	,436
	Positive	,301	,288	,244	,437	,437	,436
	Negative	-,226	-,229	-,225	-,392	-,391	-,392
Kolmogorov-Smirnov Z		4,562	4,361	3,708	6,592	6,522	6,620
Asymp. Sig. (2-tailed)		,000	,000	,000	,000	,000	,000

a Test distribution is Normal.

b Calculated from data.

Tabela 1 – Teste de Normalidade das variáveis independentes

FONTE: saída do *software* SPSS

One-Sample Kolmogorov-Smirnov Test							
		wacc	roe	Roi	margem	giro	TxReinvest.b
N		230	230	217	217	213	230
Normal Parameters (a,b)	Mean	,1584	-2,3478	,1976	,7677	1,2950	-3,4335
	Std. Deviation	,35610	37,68806	2,08328	10,10018	6,36219	62,65409
Most Extreme Differences	Absolute	,434	,488	,419	,489	,386	,453
	Positive	,434	,444	,419	,489	,386	,410
	Negative	-,410	-,488	-,358	-,432	-,320	-,453
Kolmogorov-Smirnov Z		6,580	7,406	6,176	7,208	5,637	6,863
Asymp. Sig. (2-tailed)		,000	,000	,000	,000	,000	,000

a Test distribution is Normal.

b Calculated from data.

Tabela 2 – Teste de Normalidade das variáveis independentes (continuação)

FONTE: saída do *software* SPSS

	LU												
	BETA			CRO				MARG			TXRE		
	EVA	ALAV	GIRO	KE	LOP	LIQ	EM	MD	ROE	ROI	INV	WACC	Z
EVA	1.000	0.011	0.024	0.011	0.280	0.513	0.024	0.011	0.090	0.018	-0.017	0.004	0.463
BETA		1.000	-0.015	0.999	0.030	-0.001	-0.004	1.000	0.777	-0.007	0.002	-0.015	0.097
ALAV	0.011	1.000	-0.015	0.999	0.030	-0.001	-0.004	1.000	0.777	-0.007	0.002	-0.015	0.097
GIRO	0.024	-0.015	1.000	-0.015	-0.032	-0.035	-0.012	-0.015	0.018	0.941	-0.007	0.969	-0.058
KE	0.011	0.999	-0.015	1.000	0.030	-0.001	-0.005	0.999	0.777	-0.008	0.002	-0.015	0.097
LOP	0.280	0.030	-0.032	0.030	1.000	0.859	-0.015	0.030	0.034	-0.006	0.030	-0.026	0.244
LU													
CRO													
LIQ	0.513	-0.001	-0.035	-0.001	0.859	1.000	-0.010	-0.001	0.054	-0.008	0.021	-0.033	0.340
MARG													
GEM	0.024	-0.004	-0.012	-0.005	-0.015	-0.010	1.000	-0.004	0.002	0.002	0.003	-0.001	0.125
PPL													
MD	0.011	1.000	-0.015	0.999	0.030	-0.001	-0.004	1.000	0.777	-0.007	0.002	-0.015	0.097
ROE	0.090	0.777	0.018	0.777	0.034	0.054	0.002	0.777	1.000	0.018	0.006	0.006	0.141
ROI	0.018	-0.007	0.941	-0.008	-0.006	-0.008	0.002	-0.007	0.018	1.000	0.003	0.951	-0.042

TXRE													
INV	-0.017	0.002	-0.007	0.002	0.030	0.021	0.003	0.002	0.006	0.003	1.000	0.002	0.037
WACC	0.004	-0.015	0.969	-0.015	-0.026	-0.033	-0.001	-0.015	0.006	0.951	0.002	1.000	-0.054
Z	0.463	0.097	-0.058	0.097	0.244	0.340	0.125	0.097	0.141	-0.042	0.037	-0.054	1.000

Tabela 3 – Matriz de correlação entre os pares de variáveis
 FONTE: saída do *software* SPSS

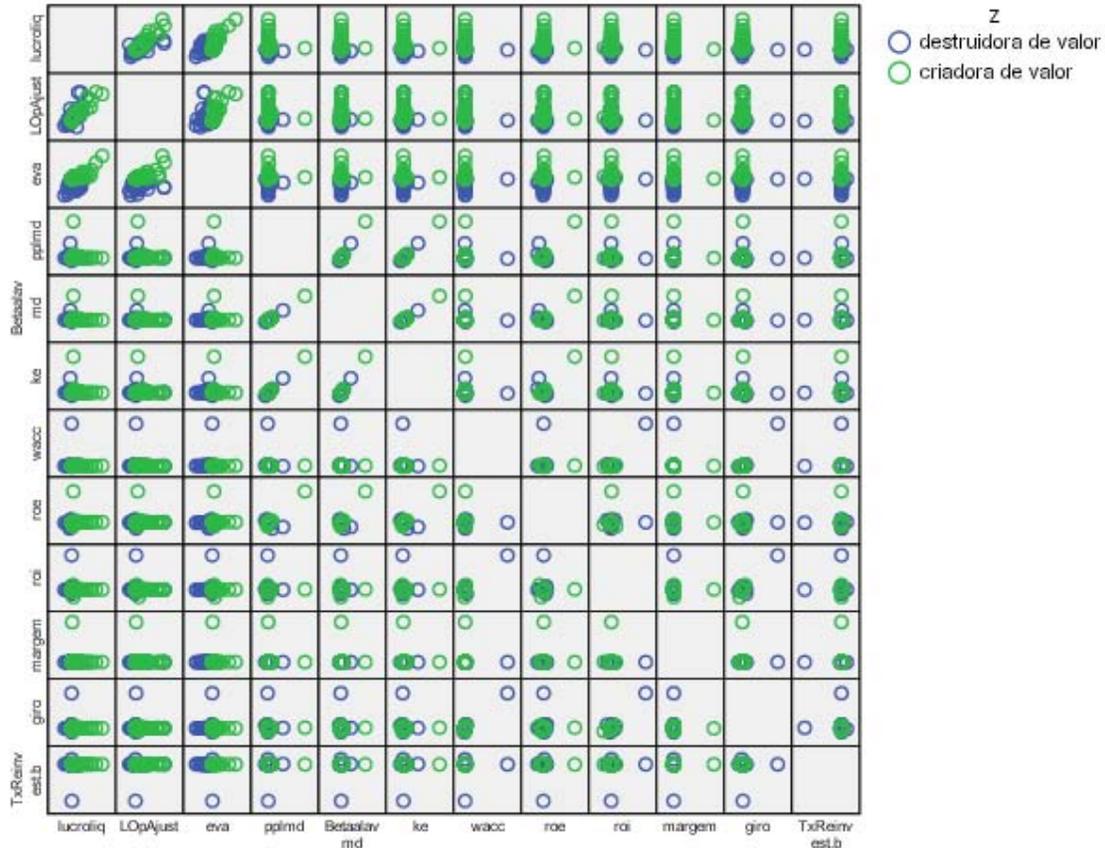


Figura 1 – Matriz com gráficos de dispersão entre os pares de variáveis
 FONTE: saída do *software* SPSS

Foram utilizadas algumas técnicas de padronização dos dados para corrigir a falta de normalidade das variáveis, mas o resultado na análise discriminante não foi significativo (as variáveis que entraram no modelo foram as mesmas e os resultados muito semelhantes). Por isso, serão analisados os dados originais mesmo com algumas condições não satisfeitas.

A análise foi feita com dois grupos: um para análise e um para validação. Os dados foram divididos em grupos de maneira randômica pelo SPSS, de modo que o grupo de validação tivesse aproximadamente 40% das observações da amostra (para replicação: utilizou-se o ponto de partida “set seed” 54321). Após a divisão dos grupos, o grupo de análise ficou com 108 observações e o grupo de validação com 84. Cerca de 10% dos casos foram descartados pelo software por apresentarem variáveis com valores perdidos (tabela 4). As tabelas apresentadas a seguir trazem os resultados encontrados.

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		122	53,0
Excluded	Missing or out-of-range group codes	0	,0
	At least one missing discriminating variable	24	10,4
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Unselected	84	36,5
	Total	108	47,0
Total		230	100,0

Tabela 4 – Descrição dos grupos criados para análise e validação

FONTE: saída do *software* SPSS

A tabela 5 demonstra quais as variáveis que podem discriminar os grupos entre criadoras e destruidoras de valor. Considerando nível de significância de 5%, as variáveis pplmd, Betaalavmd, ke, roe, roi, giro e TxReinvest.b não discriminam entre os grupos (criadoras e destruidoras de valor), pois apresentam *p-value* (Sig na tabela) superior ao nível de significância. Apenas as variáveis lucroliq (lucro líquido), LOPAdjust (lucro operacional ajustado) e margem (margem líquida) têm possibilidade de discriminar os grupos analisados.

Tests of Equality of Group Means

	Wilks' Lambda	F	df1	df2	Sig.
lucroliq	,879	16,556	1	120	,000
LOPAdjust	,933	8,626	1	120	,004
pplmd	,999	,073	1	120	,787
Betaalavmd	,999	,073	1	120	,787
ke	,999	,073	1	120	,787
wacc	,996	,463	1	120	,498
roe	,996	,494	1	120	,484
roi	,997	,376	1	120	,541
margem	,907	12,265	1	120	,001
giro	,995	,597	1	120	,441
TxReinvest.b	,998	,287	1	120	,593

Tabela 5 – Teste de igualdade entre as médias dos grupos

FONTE: Saída dada pelo *software* SPSS

A tabela 6 demonstra se existe homogeneidade entre as matrizes de covariância, um dos pressupostos para a análise de significância, como visto anteriormente. Considerando um nível de significância de 5%, deve-se rejeitar a hipótese nula, pois o *p-value* é de 0,00 ('Sig' na tabela 6), isto significa que as matrizes são heterogêneas. Então, mais um dos pressupostos da análise discriminante deixa de ser atendido. Vale ressaltar que este teste é

sensível à normalidade das variáveis. Como as variáveis não possuem distribuição normal, o resultado do teste pode estar distorcido.

Test Results		
Box's M		62,018
F	Approx.	20,100
	df1	3
	df2	41789,414
	Sig.	,000

Tests null hypothesis of equal population covariance matrices.

Tabela 6 – Resultado do teste de igualdade das matrizes de covariância

FONTE: Saída dada pelo *software* SPSS

A tabela 7 analisa se a função discriminante é significativa. Ela calcula o *p-value* para cada passo que o método *stepwise* realizou na seleção das variáveis do modelo. Como o último passo representa o modelo final proposto, percebe-se que a função discriminante é significativa para α de 5% (*p-value* < 0,05, rejeita-se a hipótese nula de que as médias dos grupos são iguais).

Wilks' Lambda				
Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,836	21,331	2	,000

Tabela 7 – Teste da Função Discriminante

FONTE: Saída dada pelo *software* SPSS

Variables Entered/Removed(a,b,c,d)							
Step	Entered	Min. D Squared					
		Statistic	Between Groups	Exact F			
				Statistic	df1	df2	Sig.
1	lucroliq	,749	destruidora de valor and criadora de valor	16,556	1	120,000	8,49E-005
2	margem	1,066	destruidora de valor and criadora de valor	11,681	2	119,000	2,33E-005

At each step, the variable that maximizes the Mahalanobis distance between the two closest groups is entered.

a Maximum number of steps is 22.

b Maximum significance of F to enter is .05.

c Minimum significance of F to remove is .10.

d F level, tolerance, or VIN insufficient for further computation.

Tabela 8 – Variáveis Selecionadas pelo Método *Stepwise*

FONTE: Saída dada pelo *software* SPSS

A tabela 8 mostra que as variáveis que entraram no modelo pelo método *stepwise* foram apenas: lucro líquido e margem líquida. Neste ponto algumas inconsistências do resultado devem ser discutidas. O primeiro resultado que chamou a atenção dos autores do trabalho foi a ausência de correlação entre variáveis altamente relacionadas na teoria como EVA[®] com ROI ou ROI com Lucro Operacional Ajustado ou ainda Lucro Líquido e

Margem Líquida. Depois, o resultado apresentado na tabela 8, também levou a questionar a capacidade de análise da função discriminante. É comprovada, até pela própria formulação do EVA[®], a influência do lucro líquido na criação de valor da empresa, bem como a influência da margem líquida (o ROI pode ser calculado pelo produto da Margem Líquida pelo Giro do Ativo), entretanto, eliminar a influência de todas as demais variáveis, como o custo de capital, e colocar na função apenas duas variáveis teoricamente muito relacionadas entre si (margem líquida é o quociente do lucro líquido pela receita líquida) é extremamente questionável.

Canonical Discriminant Function Coefficients

	Function 1
lucroliq	,000
margem	2,116
(Constant)	-,376

Unstandardized coefficients

Tabela 9 – Coeficientes da Função Discriminante

FONTE: Saída dada pelo *software* SPSS

Outro ponto questionável foi o coeficiente extremamente pequeno a ponto de aparecer ser arredondado para zero dado à variável Lucro Líquido na função. A tabela 9 apresenta os coeficientes da função discriminante sugerida. Com isso, a função discriminante deve ser assim escrita:

$$Z = -0,376 + 0,000*\text{lucroliq} + 2,116*\text{margem}$$

O valor de Z que separa o grupo das empresas criadoras de valor do grupo das empresas destruidoras de valor é chamado de $Z_{\text{crítico}}$. Ele é calculado pela média ponderada dos centróides dos grupos (criadoras e destruidoras de valor). A ponderação é feita pela quantidade de observações de cada grupo incluídas no modelo. A tabela 10 traz os valores que permitem o cálculo deste marco. Com isso, $Z_{\text{crítico}}$ é aproximadamente zero:

$$Z_{\text{crítico}} = \frac{(93 \times (-0,245)) + (29 \times 0,787)}{122} = 0,003$$

Considerando o valor encontrado do $Z_{\text{crítico}}$ (zero) e a função discriminante sugerida, é possível encontrar o valor de margem líquida a partir do qual a empresa passa a ser criadora de valor. Basta substituir o z por zero na função discriminante e encontra-se o valor da margem líquida de 0,1777. Conclui-se que apenas as empresas com margem líquida superior a 0,1777, ou 17,77%, são criadoras de valor. Esta informação seria extremamente relevante se a capacidade de classificação do modelo fosse significativa.

Functions at Group Centroids

	Function 1
z	
destruidora de valor	-,245
criadora de valor	,787

Unstandardized canonical discriminant functions evaluated at group means

Tabela 10 – Centróides dos Grupos

FONTE: Saída dada pelo *software* SPSS

A tabela 11 mostra os resultados dos testes de classificação verificando se o modelo classificou corretamente as observações. Observa-se que o poder de classificação da função não é ruim, pois no grupo de análise 78,1% dos casos foram classificados corretamente e no grupo de validação 76,4% dos casos foram classificados corretamente. É interessante observar a concentração dos erros: empresas destruidoras de valor classificadas como criadoras de valor. O inverso tem uma proporção muito pequena. Classificar a empresa destruidora de valor como criadora de valor é o erro mais prejudicial ao investidor, pois ao acreditar no potencial da empresa como o resultado da função discriminante ele está arriscado a perder todo o capital investido.

Classification Results(a,b)						
z				Predicted Group Membership		Total
				destruidora de valor	criadora de valor	
Cases Selected	Original	Count	destruidora de valor	93	4	97
			criadora de valor	24	7	31
	%		destruidora de valor	95,9	4,1	100,0
			criadora de valor	77,4	22,6	100,0
Cases Not Selected	Original	Count	destruidora de valor	60	3	63
			criadora de valor	18	8	26
	%		destruidora de valor	95,2	4,8	100,0
			criadora de valor	69,2	30,8	100,0

a 78,1% of selected original grouped cases correctly classified.

b 76,4% of unselected original grouped cases correctly classified.

Tabela 11 – Classificação dos Resultados

FONTE: Saída dada pelo *software* SPSS

Esse erro de classificação acontece porque a fórmula da função desconsidera o custo de oportunidade do capital próprio. Uma das fórmulas de cálculo do EVA[®] (ASSAF, 2003a, p. 189) é dada pela diferença entre o lucro líquido e custo de oportunidade do capital próprio ($K_e \times PL$). Se a função discriminante considerou apenas valores relacionados ao lucro líquido, ela desconsiderou o custo de oportunidade do capital próprio, ou considerou um valor único para todas as empresas do setor. Assim, é possível afirmar que as empresas destruidoras de valor classificadas incorretamente são aquelas que não conseguem remunerar adequadamente o capital do acionista, pois ao olhar o lucro (pela função discriminante) são criadoras de valor, mas ao observar o EVA[®] de fato, são destruidoras de valor (o custo de oportunidade do capital próprio “consume” todo o lucro gerado).

Considerando estes resultados, percebe-se a incapacidade da análise discriminante de indicar as variáveis que discriminam as empresas criadoras de valor das empresas destruidoras de valor. Talvez este resultado distorcido deva-se ao não cumprimento dos pressupostos do modelo estatístico apresentados. Mas, sendo assim, deve-se questionar a utilidade do modelo, pois em boa parte dos casos reais estes pressupostos não serão atendidos. Castro Júnior (2003, p.162) comparou três ferramentas para discriminação de grupos aplicando-as a empresas solventes e insolventes e concluiu:

“Comparando-se os resultados alcançados pelos modelos, foi nítida a dominância das Redes Neurais Artificiais sobre a Regressão Logística e a Análise Discriminante Linear. A ADL [Análise Discriminante Linear] foi a técnica mais pobre em termos de resultados alcançados.” (CASTRO Júnior 2003, p.162).

Observando todos esses fatores optou-se por usar também a análise por redes neurais para este conjunto de dados e comparar os resultados.

3.2 Redes Neurais

Redes neurais artificiais são sistemas de processamento de informações distribuídas, compostas por vários elementos computacionais simples que interagem através de conexões com pesos distintos. Inspiradas na arquitetura do cérebro humano, as redes neurais exibem características como a habilidade de “aprender” padrões complexos de dados e generalizar a informação aprendida (Zhang, Patuwo e Hu, 1998).

Cada elemento computacional não-linear na rede neural é chamado de nó e é densamente interconectado através de conexões diretas. Os nós operam em passos discretos, de forma análoga a uma função de dois estágios: no primeiro estágio, calcula-se a soma dos sinais de entrada, atribuindo pesos aos sinais; no segundo estágio, aplica-se, à soma dos sinais, uma função de saída, chamada de função de ativação (Hill, O'Connor e Remus, 1996).

Segundo Haykin (2001), as unidades básicas da rede são os neurônios artificiais que, por sua vez, se agrupam em camadas. Existem três categorias de camadas: a camada de entrada, a camada intermediária que pode também conter outras subcamadas e a camada de saída. Os neurônios entre as camadas são conectados por sinapses ou pesos, os quais refletem a relativa importância de cada entrada no neurônio. A camada de entrada é responsável pelas variáveis de entrada do modelo, enquanto a camada de saída contém um ou mais nós, representando os resultados finais do processamento. Nas camadas intermediárias ou camadas ocultas, o processamento é refinado, possibilitando a formação de relações não-lineares.

A Figura 2, a seguir, apresenta o modelo de um neurônio, que forma a base para o projeto de redes neurais artificiais. No modelo neural da Figura 1, podem ser destacados os seguintes componentes:

- Um conjunto de sinapses ou elos de conexão, cada sinapse é caracterizada por um peso ou força própria. Especificamente, um sinal x_j , na entrada da sinapse j conectada ao neurônio k , é multiplicado pelo peso sináptico w_{kj} . É importante notar como são escritos os índices do peso sináptico w_{kj} . O primeiro índice se refere ao neurônio em questão e o segundo ao terminal de entrada da sinapse à qual o peso se articula.
- Um somador ou integrador para somar os sinais de entrada, ponderados pelas respectivas sinapses do neurônio. As operações, neste caso específico, constituem um combinador linear. Deve-se destacar que nas redes neurais, o combinador não precisa necessariamente obedecer a uma função linear.
- Uma função de ativação para restringir a amplitude da saída de um neurônio. A função de ativação é também referida como função restritiva uma vez que restringe ou limita o intervalo permitido de amplitude do sinal de saída a um valor finito. Tipicamente, o intervalo normalizado da amplitude da saída de um neurônio é escrito como um intervalo unitário fechado $[0,1]$ ou, alternativamente, como o intervalo $[-1,1]$.

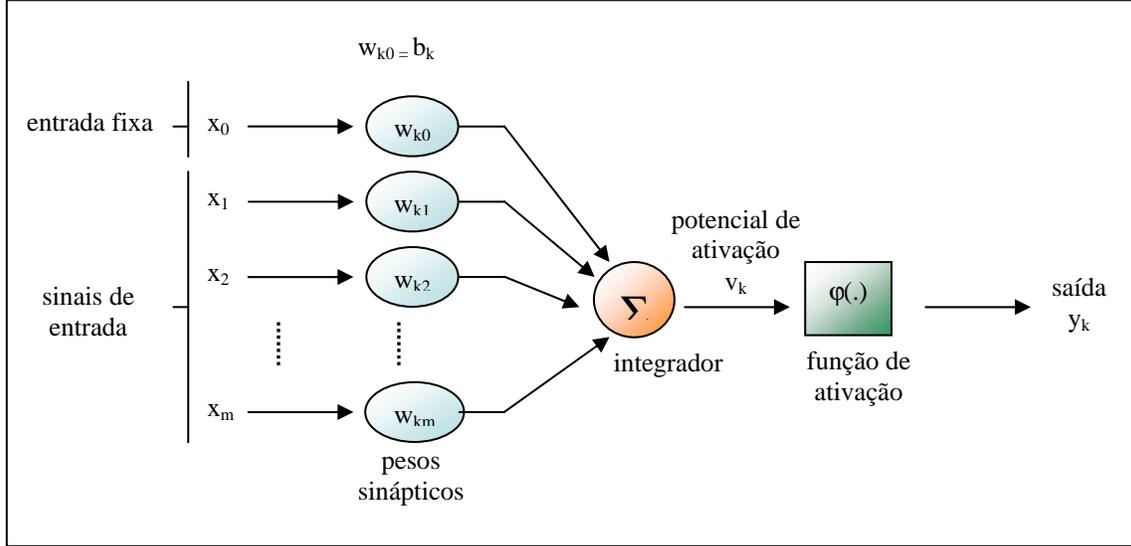


Figura 2: Modelo de um neurônio artificial

Fonte: adaptado de Haykin (2001)

O modelo neural da Figura 2 inclui também um viés aplicado externamente, representado por b_k . Este viés tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativo, respectivamente.

Em termos matemáticos, pode-se descrever um neurônio k , a partir das seguintes equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad \text{e} \quad y_k = \varphi(u_k + b_k)$$

sendo x_1, x_2, \dots, x_m os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{km}$ os pesos sinápticos do neurônio k ; u_k a saída do combinador devido aos sinais de entrada; b_k o viés; $\varphi(\cdot)$ a função de ativação e y_k o sinal de saída do neurônio.

O uso do viés b_k tem o efeito de aplicar uma transformação afim à saída u_k do combinador no modelo da Figura 2, como mostrado a seguir:

$$v_k = u_k + b_k$$

Modelos lineares têm a propriedade útil de implicar uma solução fechada para a resolução do problema de uma regressão, ou seja, de minimizar a diferença quadrática entre o_i e o valor previsto e o valor previsto p_i . Para a previsão em um período curto, o modelo linear é um ponto de partida razoável, ou um padrão reconhecido, visto que, em muitos mercados, observam-se apenas pequenas mudanças simétricas na variável a ser predita, ao longo de uma tendência de longo prazo.

No entanto, o modelo linear pode não ser preciso o bastante para aplicação em mercados financeiros voláteis, nos quais podem ocorrer processos não-lineares na evolução de dados. Movimentos lentos com aumento consistente dos preços dos ativos, seguidos por colapsos repentinos conhecidos como bolhas especulativas são comuns de acontecer. Dessa forma, o modelo linear pode falhar em capturar ou prever pontos de mudança acentuada

nos dados. Por esta razão, são utilizadas técnicas não-lineares de previsão, que podem conduzir modelos mais realistas.

Neste contexto, a rede neural é uma alternativa aos modelos de previsão tipicamente lineares e a algumas abordagens não-paramétricas para a aproximação de sistemas não-lineares. A razão para o uso de uma rede neural é simples e direta: encontrar uma abordagem ou método o qual realize bem previsões para dados gerados em processos que, freqüentemente, são desconhecidos e altamente não-lineares, com uma quantidade pequena de parâmetros e o qual seja mais fácil de estimar que os modelos não-lineares paramétricos (Oliveira, 2003).

Tendo em vista a aplicação realizada neste artigo, será discutido em mais detalhes o algoritmo de retropropagação ou *backpropagation*. Neste algoritmo, supõe-se uma rede com N neurônios na camada de entrada ($n = 1, 2, \dots, N$), J neurônios na camada intermediária ($j = 1, 2, \dots, J$) e K neurônios na camada de saída ($k = 1, 2, \dots, K$). Considera-se $W1_{jn}$ o peso entre o j -ésimo neurônio na camada intermediária e o n -ésimo neurônio da camada de entrada, e $W2_{kj}$ o peso entre o k -ésimo neurônio na camada de saída e o j -ésimo neurônio da camada intermediária, conforme ilustra Figura 3 a seguir:

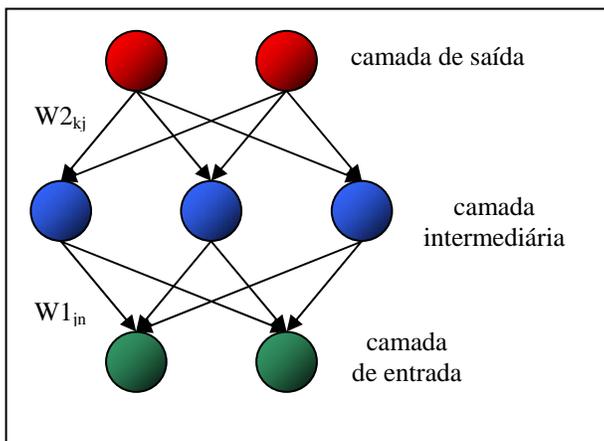


Figura 3: Estrutura da rede neural

Fonte: adaptado de Tak (1995)

A camada de entrada mostra que os nós recebem os valores dos dados de entrada. As linhas de interconexão indicam que o valor de saída de um neurônio é passado ao longo da linha até o próximo neurônio. Quando todos os dados da camada de entrada tiverem passado através da última camada, conhecida como camada de saída, um ciclo ou época terá sido realizado. Cada linha de interconexão tem um valor, chamado de peso que constitui parâmetro que opera sobre os dados associados a cada linha, através da multiplicação do valor do dado pelo peso. Estes pesos são incrementalmente ajustados durante a fase de treinamento de tal forma a alcançar o resultado de saída desejado para um determinado dado de entrada. Inicializações típicas para estes pesos encontram-se no intervalo de $[-1;1]$.

A segunda camada e todas as camadas subsequentes contêm nós de processamento, conhecidos como neurônios artificiais e são chamadas de intermediárias (ou ocultas). O processamento realizado em um neurônio artificial pode ser dividido em três passos: (i) os dados que passam, ao longo das linhas de entrada para os neurônios, são multiplicados

pelos pesos; (ii) todos esses dados que foram multiplicados pelos pesos são somados dentro do neurônio; (iii) o valor total dessa soma é passado através de uma função de transferência, cuja saída representa o valor de saída do neurônio.

A função de transferência mais simples é a função linear, adequada para um espaço linearmente separável, a qual iguala a saída à entrada. A utilização de funções de transferência não-lineares, apropriadas para um espaço não-linearmente separável, é uma das características principais da rede neural.

Usando I para indicar a camada de entrada (*input*) e O para saída (*output*), e ainda os sobrescritos i para camada de entrada, h (*hidden*) para a camada intermediária e o para camada de saída, pode-se indicar I^h como a entrada do neurônio da camada intermediária e O^i a saída do neurônio da camada de entrada.

A seguir, são apresentadas todas as etapas do algoritmo de retropropagação, de acordo com Azoff (1994):

(i) definir a função de pesos da rede, $E(w)$, conhecida como a energia ou função de custo, a qual se deseja minimizar. Considerando a metade da soma quadrática dos erros, deve-se minimizar o erro E_{total} de todos os processadores da camada de saída para todos os padrões que são apresentados à rede:

$$E_{total} = \sum_{\mu} \sum_o E(w) = \sum_{\mu} \sum_o \frac{1}{2} (T - O)^2$$

sendo, a somatória realizada sobre os padrões μ e os neurônios de saída o, T o valor alvo e O a saída observada da rede, uma função dos pesos w da rede

(ii) Calcular a saída do neurônio da camada intermediária O^h , multiplicando os pesos $W1$ pelas saídas do neurônio da camada de entrada O^i :

$$I^h = \sum_{n=1}^N W1_n(t) O^i$$

Esta entrada é o nível de ativação que determina se o neurônio pode produzir uma saída. A relação entre o nível de ativação e a saída pode ser linear ou não-linear. Considerando uma função logística e, portanto, não-linear, mas facilmente derivável, tem-se que a saída do neurônio da camada intermediária O^h é dada por:

$$O^h = \frac{1}{1 + e^{-I^h}}$$

(iii) Calcular a saída do neurônio da camada de saída O^o , multiplicando os pesos $W2$ pela saída do neurônio da camada intermediária O^h :

$$I^o = \sum_{j=1}^J W2_j(t) O^h = \sum_{j=1}^J W2_j(t) \left(\frac{1}{1 + e^{-I^h}} \right) \quad \text{e} \quad O^o = \varphi(I^o) = \frac{1}{1 + e^{-I^o}}$$

(iv) Calcular o erro:

$$E = \sum_{t=1}^n \left(T - \frac{1}{1 + e^{-I^o}} \right)$$

O mecanismo das redes neurais busca minimizar o erro total. Desta forma, as derivadas parciais do erro em relação aos pesos podem indicar o caminho que os pesos deveriam percorrer para reduzir o erro total. As derivadas parciais podem ser obtidas através do uso da regra da cadeia. Para a definição de um ajuste no peso $W2$ no algoritmo, tem-se:

$$\frac{\partial E}{\partial W2} = \frac{\partial E}{\partial I^o} \cdot \frac{\partial I^o}{\partial W2}, \text{ com}$$

$$\frac{\partial E}{\partial I^o} = \frac{\partial E}{\partial O^o} \cdot \frac{\partial O^o}{\partial I^o} = (T - O^o) \phi'(I^o) = (T - O^o) O^o (1 - O^o) = \delta^o \quad \text{e} \quad \frac{\partial I^o}{\partial W2} = O^h$$

Assim, uma variação dos pesos $\Delta W2$, para fins de implementação do algoritmo, pode ser definida como:

$$\Delta W2(t) = \alpha \frac{\partial E}{\partial W2} = \alpha \delta^o O^h$$

De modo análogo, para o ajuste no peso $W1$, tem-se:

$$\frac{\partial E}{\partial W1} = \frac{\partial E}{\partial I^h} \cdot \frac{\partial I^h}{\partial W1}, \text{ com}$$

$$\frac{\partial E}{\partial I^h} = \frac{\partial E}{\partial O^h} \cdot \frac{\partial O^h}{\partial I^h} = O^o (1 - O^o) \sum \delta^o W2 = \delta^h \quad \text{e} \quad \frac{\partial I^h}{\partial W1} = O^i$$

Com isso, o ajuste no peso pode ser realizado através de:

$$\Delta W1(t) = \alpha \frac{\partial E}{\partial W1} = \alpha \delta^h O^i$$

(v) Ajustar $W2$ e $W1$, em relação à próxima iteração:

$$W2(t+1) = W2(t) + \Delta W2(t) - \Theta (W2(t) - W2(t-1))$$

$$W1(t+1) = W1(t) + \Delta W1(t) - \Theta (W1(t) - W1(t-1))$$

onde α é a taxa de aprendizagem e $\Theta (W2(t) - W2(t-1))$ é o termo de momento.

No modelo, Θ é o coeficiente de momento que recebe um valor no intervalo $[0,1]$, usado para conferir à atualização dos pesos uma memória de sua última atualização, suavizando as forças que afetam as mudanças nos pesos.

(vi) Avaliar a recorrência do algoritmo. Se a lista de épocas foi exaurida, deve-se ir até o próximo passo, caso contrário, o algoritmo implica na apresentação do próximo conjunto de treinamento à camada de entrada e na volta ao passo inicial. Tendo completado uma época, e acumulado os erros de todos os padrões de entrada, os erros obtidos são avaliados e comparados com os níveis de tolerância.

A característica que confere à rede neural um poder de previsão é o processamento paralelo. Adicionalmente ao processamento sequencial dos sistemas lineares típicos, no qual apenas entradas observadas são utilizadas para predizer uma saída observada, pela ponderação dos neurônios de entrada, os neurônios na camada oculta processam as entradas de forma paralela, de tal forma a melhorar as predições. Esta rede de camada simples alimentada adiante, também chamada de *multi perceptron*, com uma camada escondida é a rede neural mais básica e comumente utilizada em aplicações econômicas e financeiras.

Mais genericamente, a rede representa o modo como o cérebro humano processa os dados sensoriais de entrada, recebido como neurônios de entrada, e internamente reconhecidos como um neurônio de saída. Com o desenvolvimento do cérebro, mais e mais neurônios são interconectados e mais sinapses são formadas. Os sinais de diferentes neurônios trabalham de forma paralela, em um número cada vez maior de camadas escondidas, e são combinados pelas sinapses para produzir *insights* e reações.

Nesse estudo foi utilizada uma rede do tipo *Multilayer Perceptron* (MLP), já que suas camadas intermediárias possibilitam a aproximação de funções contínuas lineares e não-lineares das entradas. O algoritmo de treinamento implementado foi o

backpropagation, o mais popular aplicado às redes MLP. Este método consiste em um aprendizado supervisionado, onde são fornecidos os valores de entrada e saída. Ele é realizado em duas etapas: a *forward* e a *backward*. Na etapa *forward*, a rede calcula os valores da saída a partir dos dados de entrada. Na fase seguinte, a *backward*, são utilizados os pesos associados a cada conexão de acordo com as diferenças entre as saídas obtidas e as desejadas, conforme já discutido anteriormente.

Para construção do modelo foi utilizada a mesma base de dados utilizada na Análise Discriminante e as mesmas variáveis independentes: Lucro Líquido, Lucro Operacional Ajustado, P/PL médio, Beta alavancado médio, Custo de capital próprio (K_e), custo médio ponderado de capital (WACC), Retorno sobre Patrimônio Líquido (ROE), Retorno sobre investimento (ROI), margem e giro. A investigação focou-se no estudo da efetivação da agregação ou destruição de valor, mais especificamente EVA[®] positivo ou negativo. A variável dependente foi considerada zero para destruição de valor e um para agregação de valor.

Ou seja, buscou-se identificar um modelo que permitisse uma análise de agregação de valor com base no conhecimento das variáveis que mais impactam no EVA[®]. Neste contexto, a amostra foi dividida em dois grupos: as empresas criadoras e as empresas destruidoras de valor. Foram consideradas empresas que criam valor aquelas que apresentavam EVA[®] positivo no referido ano.

A construção da rede neural para o cálculo do *status* de criação ou destruição de valor foi realizada através do *software Statistica da StatSoft*© 6.1. Os dados foram divididos em três bases com diferentes finalidades: treinamento, validação e teste. A base de treinamento foi composta por 50% dos dados, a de validação por 25% da amostra e a de teste por 25%. As três bases foram organizadas de modo que as proporções originais das empresas criadoras e das empresas destruidoras mantivessem praticamente as mesmas.

Em relação à arquitetura da rede, optou-se por utilizar a função de ativação sigmoideal. A taxa de aprendizagem utilizada nos treinamentos foi de 0,06. Várias redes com diferentes quantidades de neurônios nas camadas ocultas foram treinadas. O objetivo era encontrar as arquiteturas que apresentassem os melhores resultados.

Após a etapa de treinamento, o algoritmo de redes neurais foi processado, utilizando diferentes métodos. A rede que envolve a propagação em múltiplas camadas (*multiplayer propagation* ou MPL) teve desempenho superior, em relação às demais redes: rede linear, rede neural polinomial (*polynomial neural network* ou PNN) e duas redes com função radial (*radial basis function* ou RBF). Na PNN, a rede correlaciona variáveis de entrada e variáveis de meta através do uso de regressão polinomial. A RBF (Broomhead e Lowe, 1998) constitui uma alternativa ao *multilayer perceptron* e é motivada pela resposta que é ajustada localmente, como acontece nos neurônios biológicos.

A melhor rede foi considerada a MLP com média 65% de acertos na identificação das empresas que agregam ou destroem valor. Considerando as previsões feitas pela rede que apresentou melhor desempenho, a tabela 12 demonstra o comparativo das respostas.

Ano	Treinamento	Seleção	Teste	Número de Neurônios		Resultado
				Entrada	Intermediária	
1996	89,5%	62,5%	87,5%	10	7	68%
1997	88,2%	85,7%	71,4%	10	10	68%
1998	94,7%	87,5%	62,5%	10	12	79%
1999	94,4%	77,8%	75%	8	4	66%

2000	87,5%	62,5%	62,5%	9	11	63%
2001	86,7%	71,4%	57,1%	10	9	55%
2002	84,6%	50%	66,7%	6	11	55%
2003	90,9%	80%	75%	9	9	66%

Tabela 12: Desempenho da rede neural por ano

Fonte: elaborado pelos autores

Apesar de apresentar níveis de acertos razoáveis, um problema desta ferramenta é que ela não mostra quais as variáveis que discriminam os grupos. Por outro lado, as redes neurais montam um modelo que pode ser utilizado como modelo de previsão de criação de valor. Apesar do objetivo do trabalho ser identificar variáveis que discriminam os grupos de empresas conforme a criação de valor, um modelo destes teria utilidade similar para o investidor.

Para um estudo futuro, seria interessante testar a capacidade desta mesma rede neural de identificar empresas de outros setores como criadoras ou destruidoras de valor. Além de ser uma forma de testar a utilidade da rede desenvolvida neste estudo, seria uma forma de comparar a criação de valor entre diversos setores. Esta sugestão pode ser aplicada também à Análise Discriminante.

5. Conclusões

Dentre os vários objetivos de uma empresa, o consenso pela teoria das finanças é o da criação de riqueza aos acionistas. Sendo assim, maximizando o valor dos acionistas, por dedução, ela estará refletindo a utilização eficiente dos recursos econômicos de uma sociedade. A gestão baseada em valor prioriza melhorar o processo de tomada de decisões estratégicas e operacionais na organização global, partindo-se da importância na atribuição feita aos principais direcionadores de valor da entidade.

Neste contexto, o objetivo do artigo era identificar variáveis que discriminassem as empresas criadoras das empresas destruidoras de valor com base em dados empíricos. Assim, foi definido um grupo de variáveis com base em estudos sobre criação de valor econômico e buscou-se testar quais as que mais contribuíam da discriminação das empresas. Algumas características do conjunto de dados, como a falta de linearidade, impediram que a análise discriminante respondesse adequadamente ao problema. Então, optou-se por usar redes neurais com o mesmo conjunto de dados. Os resultados das redes neurais foram satisfatórios, mas a ferramenta não fornece as variáveis que mais discriminam os grupos estudados, apenas elabora um modelo que pode ser usado para novas observações.

Com isso, os resultados trazidos pelas ferramentas utilizadas não se mostraram capazes de apresentar as variáveis que melhor discriminam as empresas criadoras das empresas destruidoras de valor. Entretanto, vale ressaltar a importância de apontar características capazes de distinguir tais grupos de empresas de modo preditivo, ou seja, antes de ter acesso aos números contábeis publicados para cálculo do EVA[®].

Então, sugere-se testar no modelo de redes neurais deste trabalho observações de outros setores, com o objetivo de formar futuramente uma rede neural que classifique empresas de qualquer setor. Além disso, sugere-se também como estudo futuro tentar corrigir todos os desvios apresentados na análise discriminante, observando a manutenção

das características iniciais da amostra. E ainda, pode-se pensar em estender o estudo a outros setores.

Os estudos feitos com o conjunto de dados do setor de metalurgia e siderurgia mostram a possibilidade de se evoluir no estudo da discriminação entre empresas criadoras e empresas destruidoras de valor.

Diante de tais considerações, o estudo atingiu seus objetivos de enaltecer a teoria e a utilização da contabilidade na moderna visão de criação de valor, tornando-se indispensável e útil para os executivos e investidores. Além de ter mantido uma forte interdisciplinaridade no estudo da contabilidade e finanças. A premissa fundamental do trabalho científico além da consistência teórica é justificar aplicabilidade prática. Nesse estudo observa-se o atendimento às duas questões anteriormente ressaltadas.

6. Bibliografia

ARAUJO, Adriana M. Procópio de. **Ajustes na Contabilidade Tradicional para uma Contabilidade Baseada em Valor**. SP. Tese (Doutorado) – FEA/USP, 2002.

ASSAF NETO, Alexandre. **Finanças Corporativas e Valor**. SP: Atlas, 2003a.

_____. **Contribuição ao estudo da Avaliação de Empresas no Brasil**: uma aplicação prática. Ribeirão Preto, 2003b. Tese (Livre-Docência) – Departamento de Contabilidade, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Universidade de São Paulo.

AZOFF, E. M. **Neural network time series forecasting of financial markets**. John Wiley & Sons. 1994.

BONÍZIO, Roni C. **Análise de Sensibilidade do valor econômico agregado: um estudo aplicado nas empresas de capital aberto no Brasil**. SP. Tese (Doutorado) – FEA/USP, 2005.

BROOMHEAD, D.S.; LOWE, D. Multivariable Functional Interpolation and Adaptive Networks. *Complex Systems*, v. 2, p. 321-355, 1988.

CASTRO JR., Francisco Henrique Figueiredo de. **Previsão de Insolvência de Empresas Brasileiras Usando Análise Discriminante, Regressão Logística e Redes Neurais**. SP, Dissertação (Mestrado) – FEA/USP, 2003.

COPELAND, Tom; KOLLER, Tim; MURRIN, Jack. **Valuation**: measuring and managing the value of companies. 3rd ed. New York: John Wiley, 2000.

HAIR JR, Joseph F.; ANDERSON, Rolph E.; TATHAM, Ronald L; BLACK, William C. **Análise Multivariada de Dados**. 5^a ed. Bookman: Porto Alegre, 2005.

HAYKIN, S. **Redes Neurais: Princípios e Prática**. 2.ed. Porto Alegre: Bookman, 2001.

HILL, T.; O'CONNOR, M.; REMUS, W. *Neural Network Models for Time Series Forecasts*. **Management Science**. v. 42, n. 7, p. 1082-1092, 1996.

LIMA, Fabiano Guasti. **Um método de análise e previsão de sucessões cronológicas unidimensionais lineares e não lineares**. Faculdade de Economia, Administração e Contabilidade. São Paulo, 2004. Tese (Doutorado em Administração). Faculdade de Economia, Administração e Contabilidade: Universidade de São Paulo, 2004.

- MARSHALL, Alfred. *Principles of Economics*. New York: MacMillan & Co., 1890, v. 1.
- MAROCO, João. **Análise Estatística: Com Utilização do SPSS**. 2ª ed. Edições Silabo: Lisboa, 2003.
- MARTIN, John D.; PETTY, J. William. **Gestão Baseada em Valor: a resposta das empresas à revolução dos acionistas**. Qualitymark: RJ, 2004.
- OLIVEIRA, M. A. Previsão de sucessões cronológicas econômico-financeiras por meio de redes neurais artificiais recorrentes de tempo real e de processos ARMA-GARCH: um estudo comparativo quanto à eficiência de previsão. Faculdade de Economia, Administração e Contabilidade. São Paulo, 2003. Dissertação (Mestrado em Administração). Faculdade de Economia, Administração e Contabilidade: Universidade de São Paulo, 2003.
- RAPPAPORT, Alfred. **Gerando valor para o Acionista: um guia para administradores e investidores**. SP: Atlas, 2001.
- STERN, Joel M.; SHIELY, John S.. *The EVA Challenge: Implementing Value Added Change in an Organization*. New York: John Wiley & Sons, 2001.
- TAK, B. **A new method for forecasting stock prices using artificial neural network and ondaleta theory**. 1995. 107 f. Tese (Doutorado em Economia). Universidade de Pensilvânia, Estados Unidos, 1995.
- ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with Artificial Neural Networks: the State of the Art. *International Journal of Forecasting*, v. 14, p. 35–62, 1998.

ⁱ John Kay coloca que a análise original de valor agregado, ou lucro supranormal, é de David Ricardo (1919). Ver: KAY, John. **Fundamentos do Sucesso Empresarial**. Rio de Janeiro: Campus, 1996. p. 33.

ⁱⁱ Essa base de dados encontra-se disponibilizada nas instituições de origem dos autores para pesquisas.

ⁱⁱⁱ EVA é marca registrada da Stern Stewart & Co.