

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E CONTABILIDADE

PROGRAMA DE PÓS-GRADUAÇÃO EM ADMINISTRAÇÃO

**PREVISÃO DE INSOLVÊNCIA DE EMPRESAS BRASILEIRAS USANDO
ANÁLISE DISCRIMINANTE, REGRESSÃO LOGÍSTICA E REDES NEURAIS**

FRANCISCO HENRIQUE FIGUEIREDO DE CASTRO JUNIOR

Orientador: Prof. Dr. Ronaldo Zwicker

São Paulo

2003

FRANCISCO HENRIQUE FIGUEIREDO DE CASTRO JUNIOR

**PREVISÃO DE INSOLVÊNCIA DE EMPRESAS BRASILEIRAS USANDO
ANÁLISE DISCRIMINANTE, REGRESSÃO LOGÍSTICA E REDES NEURAIS**

Dissertação apresentada ao Departamento de Administração da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo como requisito parcial para obtenção do título de Mestre em Administração.

São Paulo

2003

FICHA CATALOGRÁFICA

Castro Junior, Francisco Henrique Figueiredo de

Previsão de insolvência de empresas brasileiras usando análise discriminante, regressão logística e redes neurais / Francisco Henrique Figueiredo de Castro Junior. -- São Paulo : FEA/USP, 2003.
169 p.

Dissertação - Mestrado

Bibliografia.

1. Administração financeira 2. Insolvência 3. Falência 4. Redes Neurais I. Faculdade de Economia, Administração e Contabilidade da USP II. Título.

CDD – 658.15

REITOR DA UNIVERSIDADE DE SÃO PAULO

Prof. Dr. Adolpho José Melfi

**DIRETORA DA FACULDADE DE ECONOMIA, ADMINISTRAÇÃO E
CONTABILIDADE**

Profa. Dra. Maria Tereza Leme Fleury

CHEFE DO DEPARTAMENTO DE ADMINISTRAÇÃO

Prof. Dr. Eduardo Pinheiro Gondin de Vasconcellos

**COORDENADOR DO PROGRAMA DE PÓS-GRADUAÇÃO EM
ADMINISTRAÇÃO**

Prof. Dr. Isak Kuglianskas

*Ao Prof. Dr. Cleber Aquino,
pela orientação e incentivo
aos meus estudos acadêmicos.*

Agradecimentos

À minha amada esposa Luciana, pelo incentivo e estímulo, principalmente nas horas mais difíceis, encorajando-me a escrever esta dissertação. Sem ela, as inúmeras horas dedicadas a este esforço teriam sido ainda mais longas. Seu exemplo de dedicação e desempenho acadêmico são também para mim um grande fator motivador.

À minha mãe, pela esmerada educação que me proporcionou, pelos exemplos de força e bravura para enfrentar situações difíceis, e pela conduta e caráter sempre pautados na retidão. À minha avó Antonieta (*in memoriam*), pelo seu amor oferecido incondicionalmente aos netos; que sempre se orgulhou de nossas conquistas e que partiu deixando saudades e boas lembranças. À minha querida irmã, que me faz orgulhoso por seu desempenho escolar e dedicação aos estudos.

Ao meu orientador Prof. Dr. Ronaldo Zwicker, pela ajuda na definição do tema de pesquisa, pelos momentos de discussão, pelo incentivo e por suas cobranças sempre muito justas.

Ao Prof. Dr. José de Oliveira Siqueira, com quem tive o privilégio de aprender e gostar ainda mais de estatística e por sua contribuição valiosa a este trabalho ao participar como membro das bancas de qualificação e defesa, oferecendo seus conhecimentos sempre atualizados.

Ao Prof. Dr. Eduardo Kazuo Kayo, por sua participação contributiva como membro da banca de defesa.

Aos professores do Departamento de Administração, em especial ao Prof. Dr. Rubens Famá, com quem tive o prazer de uma convivência acadêmica proveitosa, pelo exemplo de dedicação acadêmica, pela presença em minha banca de qualificação e por seus ricos comentários, que estão permeados neste texto.

Aos meus colegas de pós-graduação, pela convivência e aprendizado obtidos durante o programa de Mestrado, em especial aos meus amigos Lucas Ayres, Alexandre Di Miceli, Héber Pessoa e Eduardo Leite, que fazem uma brilhante carreira

acadêmica nesta Faculdade. Também agradeço aos meus amigos Humberto Halliday e Luis Eduardo Carvalho, pelo prazer de desfrutar de uma amizade que levarei para toda a vida.

Ao Grupo Marquise, por proporcionar em seu Programa de *Trainees* uma oportunidade para meu crescimento pessoal, profissional e acadêmico. Ao Sr. Newton Albuquerque, por ter sido um incentivador deste Programa e por acreditar no potencial dos jovens.

RESUMO

Estudos com o objetivo de prever insolvência de empresas e que fazem uso de técnicas estatísticas modernas são conduzidos desde a década de 1960. Esta linha de pesquisa, que inicialmente usou técnicas univariadas, e em seguida incorporou as análises multivariadas, hoje emprega largamente técnicas que fazem uso de inteligência artificial e que necessitam uma grande capacidade de processamento computacional. Esta evolução trouxe melhorias contínuas aos resultados alcançados e hoje é possível afirmar que os demonstrativos financeiros de empresas quando analisados adequadamente são uma fonte importante de informação para a previsão de insolvência. Esta pesquisa teve como principal objetivo desenvolver e comparar modelos estatísticos usando as técnicas de Análise Discriminante Linear, Regressão Logística e Redes Neurais Artificiais a fim de investigar qual delas oferece os melhores resultados. A amostra foi composta por 40 empresas, divididas em dois grupos: o primeiro com empresas formalmente insolventes segundo os critérios da legislação brasileira, e o segundo com empresas sem tais problemas. Foram usadas inicialmente 16 variáveis para predição e empregou-se um critério de seleção de variáveis baseado nos melhores subconjuntos possíveis ao invés do *stepwise*. Foi tomado especial cuidado com os pré-requisitos das técnicas, sobretudo da Análise Discriminante, como normalidade e ausência de multicolinearidade das variáveis independentes. Os resultados das previsões obtidas com os modelos foram coerentes com o esperado, ou seja, a Análise Discriminante teve um desempenho inferior à Regressão Logística que também foi superada pelas Redes Neurais Artificiais.

ABSTRACT

Researches in bankruptcy prediction of companies that make use of modern statistics techniques are being held since the 1960's. This branch of study, which initially employed univariate techniques, and then assimilated the multivariate techniques today uses artificial intelligence, a techniques that needs a great computational processing capability. This evolution brought continuing improvements to the results achieved and today is possible to say that financial statements when properly analyzed are a good source of information to the prediction of financial distress. This research aimed mainly the development of prediction models using Discriminant Analysis, Logistic Regression and Artificial Neural Networks so that they could be compared in terms of predictive capabilities. The sample consisted of 40 firms divided in 2 groups (bankrupt and non bankrupt companies) according to the Brazilian bankruptcy law. The 16 initial predictors were selected to enter the model according to the best subsets procedure in order than the stepwise procedure. Special attention was taken to accomplish the pre-requisites of the techniques, above all the Discriminant Analysis, like normality and lack of multicollinearity of the independent variables. The findings of the predictions were reasonable and according to what was expected: the Discriminant Analysis was outperformed by the Logistic Regression that was also outperformed by the Artificial Neural Networks.

SUMÁRIO

1. PROBLEMA DE PESQUISA	1
1.1. Introdução.....	1
1.2. Objetivos da Pesquisa.....	4
1.2.1 <i>Objetivos Gerais</i>	4
1.2.2 <i>Objetivos Específicos</i>	5
1.3. Justificativa da Pesquisa.....	5
1.4. Delimitação.....	6
1.5. Estrutura do Trabalho.....	6
2. REVISÃO DA LITERATURA	8
2.1. Dificuldade Financeira e Insolvência.....	8
2.2. Reestruturação de Empresas.....	12
2.3. Custos e Benefícios da Dificuldade Financeira.....	17
2.3.1 <i>Benefícios da dificuldade financeira</i>	18
2.3.2 <i>Custos da dificuldade financeira</i>	21
2.4. Modelos de Previsão de Falência.....	24
2.4.1 <i>Estudos Univariados</i>	25
2.4.1.1 Estudo de Beaver (1967).....	25
2.4.1.2 Estudo de Kanitz (1976).....	28
2.4.2 <i>Modelos Discriminantes Multivariados</i>	31
2.4.2.1 Modelo Z-Score de Altman (1968).....	31
2.4.2.2 Modelo ZETA [®] Analysis (1977).....	45
2.4.2.3 Modelo de Altman, Baidya e Dias (1979).....	48
2.4.3 <i>Modelos Usando a Regressão Logística</i>	54
2.4.3.1 Modelo de Ohlson (1980).....	54
2.4.4 <i>Uso de Redes Neurais Artificiais em Previsão de Falência</i>	59
2.4.4.1 Estudo de Odom e Sharda (1993).....	64
2.4.4.2 Estudo de Altman, Marco & Varetto (1994).....	65
2.4.4.3 Estudo de Poddig (1995).....	67
2.4.4.4 Estudo de Almeida e Dumontier (1996).....	71
3. METODOLOGIA	73
3.1. Definição Operacional das Variáveis.....	74

3.1.1	<i>Variáveis de Estrutura de Capital e Endividamento</i>	75
3.1.2	<i>Variáveis de Liquidez</i>	75
3.1.3	<i>Variáveis de Rentabilidade</i>	76
3.1.4	<i>Variáveis de Atividade e Rotação</i>	77
3.2.	Amostragem	78
3.3.	Técnicas de Análise de Dados Empregadas	79
3.3.1	<i>A Análise Discriminante Linear</i>	80
3.3.1.1	Problema de Pesquisa	82
3.3.1.2	Modelo de Pesquisa	82
3.3.1.3	Pressupostos da ADL	83
3.3.1.4	Estimação da função Discriminante	86
3.3.1.5	Interpretação dos Resultados	87
3.3.1.6	Validação dos Resultados	87
3.3.2	<i>A Regressão Logística</i>	88
3.3.2.1	Representação da Variável Dependente Binária	88
3.3.2.2	Estimação do Modelo de Regressão Logística	89
3.3.2.3	Interpretação dos Coeficientes	91
3.3.2.4	Validação dos Resultados	92
3.3.3	<i>Redes Neurais Artificiais</i>	92
3.4.	Pacotes Estatísticos Utilizados	97
4.	RESULTADOS DA PESQUISA	98
4.1.	Análise Discriminante	98
4.1.1	<i>Análise de Normalidade dos Dados</i>	98
4.1.2	<i>Seleção das Variáveis do Modelo Discriminante</i>	102
4.1.2.1	Critério R^2	104
4.1.2.2	Critério do R^2 ajustado	105
4.1.2.3	Critério C_p	106
4.1.2.4	Critério \sqrt{MSE}	107
4.1.3	<i>Análise de Multicolinearidade</i>	108
4.1.3.1	Sub-conjunto $[X_{03}, X_{05}, X_{t06}]$	109
4.1.3.2	Sub-conjunto $[X_{05}, X_{t06}]$	110
4.1.4	<i>Desenvolvimento do Modelo Discriminante</i>	111
4.1.4.1	ADL para o sub-conjunto $[X_{03}, X_{05}, X_{t06}]$	111
4.1.4.2	ADL para o sub-conjunto $[X_{05}, X_{t06}]$	121

4.2. Regressão Logística.....	129
4.2.1 <i>Seleção de Variáveis da Regressão Logística</i>	130
4.2.1.1 Critério R^2	132
4.2.1.2 Critério do R^2 ajustado	132
4.2.1.3 Critério C_p	133
4.2.1.4 Critério \sqrt{MSE}	134
4.2.2 <i>Análise de Multicolinearidade</i>	135
4.2.2.1 Sub-conjunto $[X_1; X_2; X_3; X_8; X_{15}; X_{r6}]$	135
4.2.2.2 Sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$	136
4.2.2.3 Sub-conjunto $[X_1; X_2; X_3; X_8; X_9; X_{11}; X_{12}; X_{15}; X_{r6}]$	137
4.2.2.4 Sub-conjunto $[X_3; X_{14}; X_{r6}]$	138
4.2.3 <i>Desenvolvimento do Modelo Logístico</i>	138
4.2.3.1 Modelo Logístico para o sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$.	139
4.2.3.2 Modelo Logístico para o sub-conjunto $[X_3; X_{14}; X_{r6}]$	141
4.3. Redes Neurais Artificiais.....	142
4.3.1 <i>Rede com Treinamento de 80% dos Casos</i>	143
4.3.2 <i>Rede com Treinamento de 60% dos Casos</i>	150
4.3.3 <i>Rede com Treinamento de 50% dos Casos</i>	155
5. CONCLUSÕES E RECOMENDAÇÕES	160
6. REFERÊNCIAS BIBLIOGRÁFICAS	165

LISTA DE FIGURAS

FIGURA 1: INSOLVÊNCIA ASSOCIADA A SALDOS E INSOLVÊNCIA ASSOCIADA A FLUXOS. ADAPTADO DE: ROSS, WESTERFIELD & JAFFE (2002).....	10
FIGURA 2: INSOLVÊNCIA ASSOCIADA A SALDOS (FIGURA A) E INSOLVÊNCIA ASSOCIADA A FLUXOS (FIGURA B): DIFERENCIAÇÃO NA LINHA DO TEMPO. FIGURA ADAPTADA DE WRUCK (1990, P.423).	11
FIGURA 3: EXEMPLO DA CLASSIFICAÇÃO DICOTÔMICA FEITA POR BEAVER (1967).....	27
FIGURA 4: PROCESSADORES DAS REDES NEURAIS ARTIFICIAIS.....	60
FIGURA 5: REPRESENTAÇÃO UNIVARIADA DOS ESCORES DISCRIMINANTES.....	81
FIGURA 6: EXEMPLO DE REGRESSÃO LOGÍSTICA.	89
FIGURA 7: CURVAS DE REGRESSÃO LOGÍSTICA, RESPECTIVAMENTE BEM AJUSTADA E MAL AJUSTADA.	90
FIGURA 8: ARQUITETURA DA REDE A SER UTILIZADA NESTE ESTUDO.	94

LISTA DE GRÁFICOS

GRÁFICO 1: PERCENTUAL DA ALTA ADMINISTRAÇÃO ORIGINAL QUE PERMANECE NA FIRMA APÓS UM PERÍODO DE DIFICULDADE FINANCEIRA CONFORME O MÉTODOS DE REESTRUTURAÇÃO FINANCEIRA. FONTE: GILSON (1991, p.651).....	20
GRÁFICO 2: ANÁLISE DISCRIMINANTE LINEAR. EXEMPLO EXTRAÍDO DE ALTMAN (1983, p.104). NO GRÁFICO, AS EMPRESAS SAUDÁVEIS SÃO DENOTADAS COM UM X, E AS EMPRESAS NÃO SAUDÁVEIS COM UM O. AS MÉDIAS DOS GRUPOS ESTÃO CIRCULADAS, E A LINHA AB É A QUE MELHOR DISCRIMINA OS DOIS GRUPOS.	34
GRÁFICO 3: TRANSFORMAÇÃO DE BOX-COX PARA A VARIÁVEL X_{04}	100
GRÁFICO 4: TRANSFORMAÇÃO DE BOX-COX PARA A VARIÁVEL X_{06}	101
GRÁFICO 5: GRÁFICO DE R^2 CONTRA O NÚMERO DE PARÂMETROS.	104
GRÁFICO 6: GRÁFICO DE R^2 AJUSTADO CONTRA O NÚMERO DE PARÂMETROS.	105
GRÁFICO 7: GRÁFICO DE C_p CONTRA O NÚMERO DE PARÂMETROS.	107
GRÁFICO 8: GRÁFICO DE \sqrt{MSE} CONTRA O NÚMERO DE PARÂMETROS.	108
GRÁFICO 9: GRÁFICO DE R^2 AJUSTADO CONTRA O NÚMERO DE PARÂMETROS.	133
GRÁFICO 10: GRÁFICO DE C_p CONTRA O NÚMERO DE PARÂMETROS.	134
GRÁFICO 11: MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO AO LONGO DOS CICLOS DE APRENDIZAGEM DA RNA-80.	144
GRÁFICO 12: MÉDIA DOS MSEs MÍNIMOS PARA DIFERENTES QUANTIDADES DE NEURÔNIOS NA CAMADA INTERMEDIÁRIA.	146
GRÁFICO 13: ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DO MODELO RNA-80.	148
GRÁFICO 14: MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO AO LONGO DOS CICLOS DE APRENDIZAGEM DA RNA-60.	150
GRÁFICO 15: MÉDIA DOS MSEs MÍNIMOS PARA DIFERENTES QUANTIDADES DE NEURÔNIOS NA CAMADA INTERMEDIÁRIA.	153
GRÁFICO 16: ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DO MODELO RNA-60.	153
GRÁFICO 17: MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO AO LONGO DOS CICLOS DE APRENDIZAGEM DA RNA-50.	155
GRÁFICO 18: MÉDIA DOS MSEs MÍNIMOS PARA DIFERENTES QUANTIDADES DE NEURÔNIOS NA CAMADA INTERMEDIÁRIA.	157
GRÁFICO 19: ANÁLISE DE SENSIBILIDADE DAS VARIÁVEIS DO MODELO RNA-50.	158

LISTA DE TABELAS

TABELA 1: MÉDIA DAS VARIÁVEIS E TESTE DE SIGNIFICÂNCIA. FONTE: ALTMAN, 1968.	39
TABELA 2: RESULTADO DO ESTÁGIO CLASSIFICATÓRIO DO MODELO DISCRIMINANTE DE ALTMAN MEDIDO UM ANO ANTES DA FALÊNCIA. FONTE: ALTMAN, 1968.	40
TABELA 3: RESULTADO DO ESTÁGIO CLASSIFICATÓRIO DO MODELO DISCRIMINANTE DE ALTMAN MEDIDO DOIS ANOS ANTES DA FALÊNCIA. FONTE: ALTMAN, 1968.	40
TABELA 4: PRECISÃO DOS TESTES DE VALIDAÇÃO PARA AS SUB-AMOSTRAS. FONTE: ALTMAN, 1968.....	42
TABELA 5: PRECISÃO DOS TESTES DE VALIDAÇÃO PARA NOVA AMOSTRA DE EMPRESAS INSOLVENTES. FONTE: ALTMAN, 1968.....	42
TABELA 6: PRECISÃO DOS TESTES DE VALIDAÇÃO PARA NOVA AMOSTRA DE EMPRESAS SOLVENTES, MAS QUE REPORTARAM PREJUÍZOS NOS ANOS DE 1958 E 1961. FONTE: ALTMAN, 1968.....	43
TABELA 7: PRECISÃO DOS TESTES DE CLASSIFICAÇÃO PARA AS EMPRESAS DA AMOSTRA ORIGINAL NO LONGO PRAZO. FONTE: ALTMAN, 1968.....	44
TABELA 8: QUADRO COMPARATIVO ENTRE A PRECISÃO DE PREVISÃO DO MODELO ZETA TM E O MODELO Z-SCORE EM PERCENTAGENS. FONTE: ADAPTADO DE ALTMAN (1983, p.137).	48
TABELA 9: PRECISÃO DOS TESTES DE CLASSIFICAÇÃO PARA EMPRESAS INSOLVENTES.	52
TABELA 10: PRECISÃO DOS TESTES DE CLASSIFICAÇÃO PARA EMPRESAS SOLVENTES.....	52
TABELA 11: RESULTADOS DOS EXPERIMENTOS DE OHLSON, 1980.....	57
TABELA 12: RESULTADOS DOS MODELOS DE REGRESSÃO LOGÍSTICA DE OHLSON, 1980.	57
TABELA 13: COMPARAÇÃO DOS RESULTADOS OBTIDOS COM A ANÁLISE DISCRIMINANTE E REDES NEURAIS. FONTE: ODOM & SHARDA (1993, pp.182).	65
TABELA 14: RESULTADOS DA ADL PARA O ESTUDO DE Poddig (1995).....	68
TABELA 15: RESULTADO DO MODELO DE RNA DE Poddig (1995) USANDO AJUSTAMENTO DE PESOS PELO CUSTO E PADRONIZAÇÃO EM CONJUNTO COM TRANSFORMAÇÃO NÃO-LINEAR DAS VARIÁVEIS.....	70
TABELA 16: RESULTADO DO MODELO DE RNA DE Poddig (1995) USANDO AJUSTE ESTATÍSTICO DOS PESOS E OTIMIZAÇÃO MANUAL DA REDE.	70
TABELA 17: TESTE DE NORMALIDADE PARA AS POSSÍVEIS VARIÁVEIS INDEPENDENTES.....	99

TABELA 18: TESTE DE NORMALIDADE PARA AS VARIÁVEIS X_{04} E X_{06} TRANSFORMADAS.	102
TABELA 19: VALORES DE R^2 , R^2 AJUSTADO, Cp E S PARA OS SUB-CONJUNTOS DO MODELO.	103
TABELA 20: ESTATÍSTICAS DE COLINEARIDADE.	109
TABELA 21: EIGENVALUES, ÍNDICE DE CONDIÇÃO E PROPORÇÕES DE VARIÂNCIA.	109
TABELA 22: ESTATÍSTICAS DE COLINEARIDADE.	110
TABELA 23: EIGENVALUES, ÍNDICE DE CONDIÇÃO E PROPORÇÕES DE VARIÂNCIA.	110
TABELA 24: ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS DO MODELO.	112
TABELA 25: TESTE DE IGUALDADE DE MÉDIA PARA AS VARIÁVEIS DO MODELO.	112
TABELA 26: TESTE M DE BOX PARA IGUALDADE DAS MATRIZES DE VARIÂNCIA-COVARIÂNCIA.	112
TABELA 27: FUNÇÃO DISCRIMINANTE CANÔNICA.	113
TABELA 28: SIGNIFICÂNCIA DA FUNÇÃO DISCRIMINANTE CANÔNICA.	113
TABELA 29: COEFICIENTES DA FUNÇÃO DISCRIMINANTE CANÔNICA.	113
TABELA 30: CARGAS DA FUNÇÃO DISCRIMINANTE.	114
TABELA 31: COEFICIENTES DAS FUNÇÕES DISCRIMINANTES LINEARES DE FISHER.	114
TABELA 32: CENTRÓIDES DAS FUNÇÕES DISCRIMINANTES CANÔNICAS.	115
TABELA 33: MATRIZ DE CLASSIFICAÇÃO FINAL PARA O GRUPO ORIGINAL E PARA A VALIDAÇÃO.	116
TABELA 34: PREDIÇÕES INDIVIDUAIS DOS CASOS PARA A AMOSTRA ORIGINAL.	118
TABELA 35: PREDIÇÕES INDIVIDUAIS DOS CASOS PARA A AMOSTRA DE VALIDAÇÃO.	119
TABELA 36: DIFERENÇA DE MÉDIA PARA AS VARIÁVEIS INDEPENDENTES ENTRE EMPRESAS CLASSIFICADAS CORRETAMENTE E INCORRETAMENTE PARA O GRUPO DE EMPRESAS INSOLVENTES.	120
TABELA 37: DIFERENÇA DE MÉDIA PARA AS VARIÁVEIS INDEPENDENTES ENTRE EMPRESAS CLASSIFICADAS CORRETAMENTE E INCORRETAMENTE PARA O GRUPO DE EMPRESAS SOLVENTES.	121
TABELA 38: TESTE M DE BOX PARA IGUALDADE DAS MATRIZES DE VARIÂNCIA-COVARIÂNCIA.	122
TABELA 39: FUNÇÃO DISCRIMINANTE CANÔNICA.	122
TABELA 40: SIGNIFICÂNCIA DA FUNÇÃO DISCRIMINANTE CANÔNICA.	122
TABELA 41: COEFICIENTES DA FUNÇÃO DISCRIMINANTE CANÔNICA.	123
TABELA 42: CARGAS DA FUNÇÃO DISCRIMINANTE.	123
TABELA 43: COEFICIENTES DAS FUNÇÕES DISCRIMINANTES LINEARES DE FISHER.	124
TABELA 44: CENTRÓIDES DAS FUNÇÕES DISCRIMINANTES CANÔNICAS.	124

TABELA 45: MATRIZ DE CLASSIFICAÇÃO FINAL PARA O GRUPO ORIGINAL E PARA A VALIDAÇÃO.	125
TABELA 46: PREDIÇÕES INDIVIDUAIS DOS CASOS PARA A AMOSTRA ORIGINAL.	126
TABELA 47: PREDIÇÕES INDIVIDUAIS DOS CASOS PARA A AMOSTRA DE VALIDAÇÃO.	127
TABELA 48: DIFERENÇA DE MÉDIA PARA AS VARIÁVEIS INDEPENDENTES ENTRE EMPRESAS CLASSIFICADAS CORRETAMENTE E INCORRETAMENTE PARA O GRUPO DE EMPRESAS INSOLVENTES.	129
TABELA 49: VALORES DE R^2 , R^2 AJUSTADO, Cp e S PARA OS SUB-CONJUNTOS DO MODELO LOGÍSTICO.	131
TABELA 50: ESTATÍSTICAS DE COLINEARIDADE.....	135
TABELA 51: EIGENVALUES, ÍNDICE DE CONDIÇÃO E PROPORÇÕES DE VARIÂNCIA.	136
TABELA 52: ESTATÍSTICAS DE COLINEARIDADE.....	136
TABELA 53: EIGENVALUES, ÍNDICE DE CONDIÇÃO E PROPORÇÕES DE VARIÂNCIA.	137
TABELA 54: ESTATÍSTICAS DE COLINEARIDADE.....	137
TABELA 55: ESTATÍSTICAS DE COLINEARIDADE.....	138
TABELA 56: EIGENVALUES, ÍNDICE DE CONDIÇÃO E PROPORÇÕES DE VARIÂNCIA.	138
TABELA 57: RESULTADOS DOS MODELOS LOGÍSTICOS USANDO O SUB-CONJUNTO [X_3 ; X_8 ; X_{15} ; X_{R6}].	140
TABELA 58: COEFICIENTES DOS MODELOS DE REGRESSÃO PARA O SUB-CONJUNTO [X_3 ; X_8 ; X_{15} ; X_{R6}].	141
TABELA 59: RESULTADOS DOS MODELOS LOGÍSTICOS USANDO O SUB-CONJUNTO [X_3 ; X_{14} ; X_{R6}].	141
TABELA 60: COEFICIENTES DOS MODELOS DE REGRESSÃO PARA O SUB-CONJUNTO [X_3 ; X_{14} ; X_{R6}].	142
TABELA 61: VALORES MÍNIMOS E DESVIO PADRÃO PARA O MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO DA RNA-80.....	145
TABELA 62: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-80 CONTENDO AS 18 VARIÁVEIS DE PREDIÇÃO.....	146
TABELA 63: VALORES MÍNIMOS E DESVIO PADRÃO PARA O MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO DA REDE COM 5 NEURÔNIOS E 5 VARIÁVEIS DE ENTRADA.....	148
TABELA 64: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-80 CONTENDO AS 5 VARIÁVEIS DE PREDIÇÃO.....	149

TABELA 65: VALORES MÍNIMOS E DESVIO PADRÃO PARA O MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO DA RNA-60.....	151
TABELA 66: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-60 CONTENDO AS 18 VARIÁVEIS DE PREDIÇÃO.....	152
TABELA 67: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-60 CONTENDO 7 VARIÁVEIS DE PREDIÇÃO.....	154
TABELA 68: VALORES MÍNIMOS E DESVIO PADRÃO PARA O MSE MÉDIO DAS 10 RODADAS PARA AS AMOSTRAS DE TREINAMENTO E VALIDAÇÃO DA RNA-50.....	156
TABELA 69: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-50 CONTENDO AS 18 VARIÁVEIS DE PREDIÇÃO.....	156
TABELA 70: RESULTADOS DE CLASSIFICAÇÃO PARA A RNA-50 CONTENDO 7 VARIÁVEIS DE PREDIÇÃO.....	159

1. Problema de Pesquisa

1.1. Introdução

As questões relativas às dificuldades financeiras de empresas têm grande importância para um público diverso como instituições financeiras fornecedoras de crédito, especialistas em investimentos, investidores em geral, auditores contábeis, consultores, gestores de empresas e empregados, acadêmicos, pesquisadores e estudantes de escolas de negócio.

O primeiro trabalho de que se tem notícia sobre este tema é o de Fitzpatrick em 1932. Em sua pesquisa, Fitzpatrick comparou dezenove empresas bem sucedidas com dezenove empresas que foram à falência durante os anos de 1920 a 1929 nos Estados Unidos. Naquela época, devido à falta de técnicas adequadas de comparação estatística, foram utilizados dois métodos: o primeiro consistia em observar alguns indicadores de desempenho da empresa e classificá-los como acima ou abaixo de um determinado padrão ideal; o outro consistia em observar a tendência desses mesmos indicadores ao longo do tempo.

Após comparar os indicadores financeiros dos dois grupos de empresas, Fitzpatrick concluiu que os índices extraídos de demonstrações contábeis poderiam fornecer indicações importantes quanto ao risco de inadimplência de uma empresa. No entanto, ele alerta que

It should be remembered, moreover, that the author does not advocate that ratios are absolute indicators of impending financial difficulties. They are, indeed,

helpful clues to be made the subject for further investigation. They are, also, tests of business health. (1932: p.21 apud Robinson, 2002: p.12).

Muito se avançou com relação às técnicas de previsão de inadimplência. Para conseguir mensurar as reais chances de uma empresa estar caminhando para uma situação de dificuldades financeiras e com risco de inadimplência e/ou concordata/falência, os pesquisadores desta área em Finanças passaram a utilizar modelos estatísticos que proporcionassem uma avaliação com precisão. Os primeiros estudos usaram a técnica da Análise Discriminante Linear (ADL): primeiramente a Análise Discriminante Univariada com o estudo de Beaver em 1966 e posteriormente a Análise Discriminante Multivariada com o de Altman em 1968. Os bons resultados alcançados após o uso dessas técnicas trouxeram maior credibilidade aos estudos de previsão de dificuldades financeiras, e serviram para difundir a realização de testes empíricos com empresas de todo o mundo. Todavia, os modelos que se basearam nessa técnica de previsão de falência foram também alvos de muitas críticas, pois tais modelos geralmente não atendiam aos rígidos pressupostos que a técnica requisitava e seus resultados embutiam um alto risco de levar à tomada de decisão errada. Para se poder aplicar a ADL adequadamente é necessário garantir requisitos rígidos como a normalidade multivariada das variáveis independentes, além da igualdade das matrizes de variância-covariância entre os grupos que representam as empresas em dificuldades financeiras e as empresas saudáveis.

Apesar de toda essa rigidez quanto aos pressupostos da técnica, foi a ADL desenvolvida por Altman (1968) que se tornou o modelo mais conhecido e utilizado entre aqueles interessados em realizar previsões dessa natureza. No Brasil, Altman realizou um trabalho em 1979 em conjunto com dois professores brasileiros onde

examinaram a experiência de empresas locais para a previsão de dificuldades financeiras (ALTMAN, BAYDIA e DIAS, 1979). Eles desenvolveram um modelo específico para classificar e prever problemas financeiros em empresas no Brasil. Este modelo, desenvolvido a partir da mesma técnica utilizada para o modelo com as empresas americanas e o modelo desenvolvido por Stephen Kanitz em 1976 foram os primeiros a levar em consideração a realidade dos dados de empresas no Brasil.

Em 1980, Ohlson inseriu a Regressão Logística (RL) na avaliação de riscos de dificuldades financeiras em empresas. A RL passou a ser então preferida pelos pesquisadores do assunto por não depender rigidamente dos pressupostos da ADL, o que tornou sua aplicabilidade muito mais ampla em diversas pesquisas realizadas. Além disso, a RL é similar a uma regressão múltipla onde as variáveis independentes são métricas e a variável dependente é não métrica (nominal).

O advento de novas tecnologias e a velocidade cada vez maior para processar dados trouxe ferramentas muito mais poderosas que passaram a ser adotadas como soluções em problemas na área de Finanças. As Redes Neurais Artificiais (RNA) são um caso específico, onde na área de análise de crédito oferece resultados alternativos aos então produzidos pela ADL e RL. As RNA reproduzem de forma simplificada o funcionamento do cérebro humano, sendo capazes de tratar dados incompletos e distorcidos produzindo resultados satisfatórios através de generalizações (Almeida e Dumontier, 1996).

Neste cenário de diferentes técnicas possíveis de serem empregadas para avaliar o caso de previsão de insolvência em empresas, este estudo busca comparar a

qualidade de predição para cada uma das possibilidades apresentadas utilizando-se de dados de empresas que atuam no Brasil.

1.2. Objetivos da Pesquisa

1.2.1 Objetivos Gerais

Como observado anteriormente por Back *et alii* (1996), duas grandes correntes são verificadas em estudos que tratam do tema previsão de dificuldades financeiras. A primeira delas é procurar por índices contábeis que possuam boa capacidade de discriminação entre as empresas em boa situação financeira e empresas que passam por dificuldades financeiras; a segunda corrente se concentra na busca e avaliação de técnicas estatísticas que melhorem cada vez mais as capacidades de previsão. A pesquisa proposta pretende atingir três grandes objetivos gerais:

- Desenvolver um novo modelo de previsão de falência para empresas que atuam no Brasil usando a Análise Discriminante Linear (ADL);
- Desenvolver um novo modelo de previsão de falência para empresas que atuam no Brasil usando a Regressão Logística (RL);
- Desenvolver um novo modelo de previsão de falência para empresas que atuam no Brasil usando Redes Neurais Artificiais (RNA).

1.2.2 *Objetivos Específicos*

- Selecionar quais indicadores são mais úteis no processo de discriminação entre empresas dos dois grupos (solventes e insolventes) para cada uma das três técnicas a serem empregadas;
- Construir uma equação usando a ADL para discriminar adequadamente empresas em situação de insolvência das empresas solventes;
- Construir uma equação usando a RL para discriminar adequadamente empresas em situação de insolvência das empresas solventes;
- Treinar uma RNA a fim de construir um modelo de rede capaz de discriminar adequadamente empresas em situação de insolvência das empresas solventes.
- Comparar a qualidade¹ de discriminação dos três modelos construídos.

1.3. Justificativa da Pesquisa

Como afirmaram Sanvicente e Minardi (1998), previsão de concordatas é um aspecto essencial de modelos de gerenciamento de risco de crédito. No Brasil especialmente, são poucos os estudos realizados nesta área, seja por falta de dados adequados a um estudo dessa natureza, seja pela crença de que indicadores contábeis não oferecem grande confiabilidade para a análise de risco de crédito de empresas. Apesar do que se possa afirmar contra a confiabilidade de dados contábeis para estudos dessa natureza em vista de escândalos de fraude contábil ou suspeita de

¹ Entende-se por qualidade a capacidade da técnica em realizar previsões *a priori* a respeito do estado da empresa quanto à sua saúde financeira.

maquiagem de balanços, os estudos em geral mostram que se analisados adequadamente, como afirmam Matias e Siqueira (1996), os balanços podem se constituir em fonte de dados fundamentais para modelos de previsão de insolvência.

Os custos envolvidos em uma eventual falência de uma empresa provocam grandes perdas a uma série de pessoas como acionistas, credores e empregados de uma firma. Boas técnicas de previsão de insolvência ajudam tanto aos gestores a serem capazes de controlar o desempenho empresarial de sua firma antes que ela ultrapasse uma situação irrecuperável quanto aos credores que investem naquele negócio e pretendem minimizar os riscos de seus investimentos.

1.4. Delimitação

São objetos desta pesquisa as empresas não financeiras de capital aberto, com atuação no Brasil entre os anos de 1996 e 2002, constantes no banco de dados da Comissão de Valores Mobiliários (CVM) e da Bolsa de Valores do Estado de São Paulo (BOVESPA).

1.5. Estrutura do Trabalho

Esta dissertação está estruturada sob a forma de cinco capítulos. Após este capítulo inicial de introdução, o Capítulo 2 apresenta a revisão de literatura pertinente ao estudo e que será a base para o entendimento do restante do texto. Em seguida, o Capítulo 3 apresenta a metodologia empregada nesta pesquisa. Os resultados e a sua discussão fazem parte do Capítulo 4. No Capítulo 5 são

apresentadas as conclusões advindas do estudo, suas limitações e recomendações para o desenvolvimento de pesquisas posteriores.

2. Revisão da Literatura

2.1. Dificuldade Financeira e Insolvência

Os termos dificuldade financeira e insolvência são muitas vezes encontrados na literatura sendo usados de forma indiscriminada. No entanto, alguns autores distinguem claramente essas duas formas de situação indesejável em empresas. Para Wruck (1990, p.421), uma empresa encontra-se em dificuldade financeira quando seu fluxo de caixa é insuficiente para cobrir as obrigações correntes. De forma semelhante, Ross, Westerfield & Jaffe (2002, p.683) definem dificuldade financeira como uma situação em que uma empresa não gera fluxo de caixa suficiente para fazer um pagamento contratualmente devido.

Em 1983, Altman (1983, p.6) comenta também o conceito de *failure*, segundo o qual a empresa se encontra em uma situação em que a taxa de retorno sobre o capital investido é significativamente e continuamente menor que as taxas mais comuns existentes para investimentos similares. Essa situação não implica que a empresa venha a estar ou esteja descumprindo com suas obrigações legais. Por outro lado, a insolvência está diretamente relacionada ao fato dela estar sem meios para realizar o pagamento de suas dívidas.

A insolvência pode ser separada em duas formas conceitualmente distintas:

- Insolvência baseada em fluxos (*flow-based insolvency*) ou insolvência técnica (*technical insolvency*), e
- Insolvência baseada em saldos (*stock-based insolvency*).

Segundo Altman (1983, p.6), a insolvência técnica diz respeito a uma falta de liquidez da firma que impossibilita o cumprimento de uma obrigação. Sendo assim, a insolvência técnica pode ser uma condição temporária. Já a insolvência baseada em saldos, segundo Wruck (1990, p.421), ocorre quando a firma tem valor econômico negativo, ou seja, o valor presente dos fluxos de caixa futuros é menor que o total de suas obrigações. Esse conceito de insolvência é mais grave, e segundo Altman

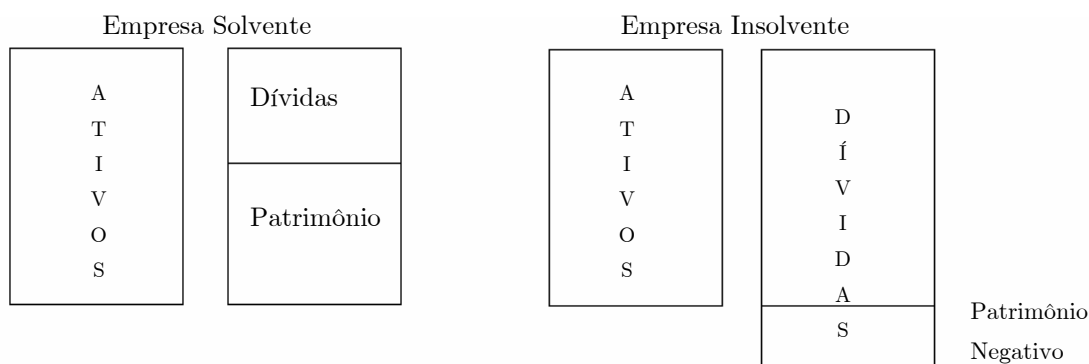
(...) indica uma situação crônica e não temporária. A firma está em uma situação em que seus passivos totais excedem uma valoração financeira justa do total de seus ativos. (1983, p.6)

A figura 1 representa esquematicamente as diferenças entre os conceitos de insolvência baseada em fluxo (ou insolvência técnica) e insolvência baseada em saldos. A insolvência associada a saldos ocorre quando o valor dos ativos de uma empresa é inferior ao valor de suas dívidas. Isso significa que o patrimônio líquido é negativo. A insolvência associada a fluxos se dá quando os fluxos de caixa da empresa são insuficientes para cobrir pagamentos determinados por contrato (Ross, Westerfield & Jaffe, 2002).

Um problema sério a ser contornado em casos de inadimplência é a assimetria de informações entre credores e acionistas. A esses últimos interessa mostrar aos credores que a empresa está passando por uma situação de insolvência baseada em fluxo, ou seja, a empresa está momentaneamente inadimplente, mas tem todas as condições de se reerguer a partir de seus fluxos de caixa futuros. Os credores, por outro lado, temem o agravamento da situação, pois não sabem o que realmente está se passando com a empresa internamente. Eles tendem a pensar que a empresa está

passando por uma situação de insolvência baseada em saldos, pois aumentaria a possibilidade deles serem recompensados com o patrimônio da firma em uma liquidação.

1. Insolvência associada a saldos



2. Insolvência associada a fluxos

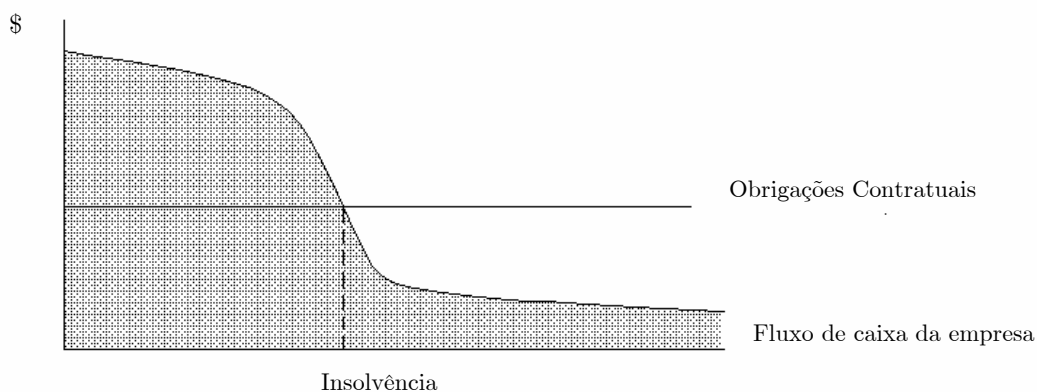


Figura 1: Insolvência associada a saldos e insolvência associada a fluxos. Adaptado de: Ross, Westerfield & Jaffe (2002).

A figura 2 apresenta as duas situações em que a empresa pode se encontrar devido a insolvências baseada em fluxos e baseada em saldos. Para resolver adequadamente o problema da assimetria de informações entre credores e acionistas, informações importantes quanto ao desempenho da empresa devem ser compartilhadas abertamente. Como esse processo de abertura de informações

(*disclosure*) nem sempre tem se mostrado eficaz, tentativas também tem sido feitas quanto à previsão de falência de empresas que passam por dificuldades financeiras. Os modelos de previsão de falência serão tratados posteriormente, na seção 4 deste capítulo. Na figura 2, as linhas horizontais representam as obrigações por unidade de tempo. O período até o segundo corte temporal é exatamente idêntico para as duas figuras A e B. Só a partir do segundo corte as empresas diferenciam-se consideravelmente. A empresa da figura A se recupera. A empresa da figura B não se recupera.

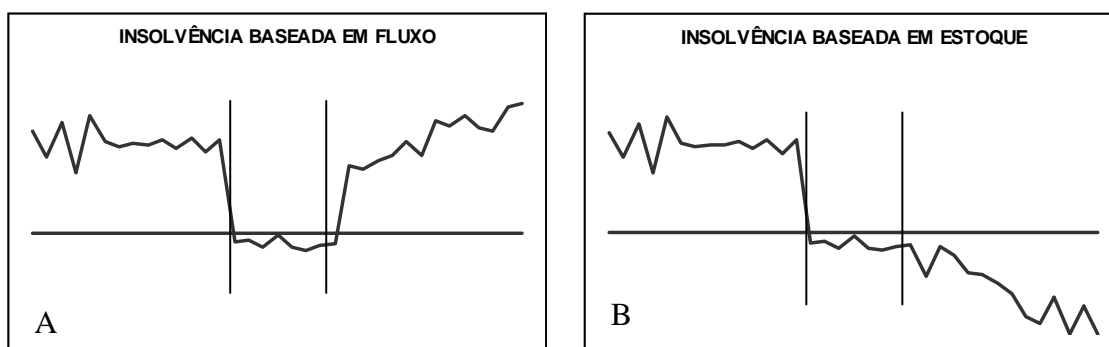


Figura 2: *Insolvência associada a saldos (Figura A) e insolvência associada a fluxos (Figura B): diferenciação na linha do tempo. Figura adaptada de Wruck (1990, p.423).*

Segundo Wruck (1990, pp. 421-2), a insolvência baseada em fluxo dá aos credores prejudicados o direito de demandar uma reestruturação porque o contrato com a firma foi violado. Se no entanto a firma está insolvente baseado em saldos, mas vem cumprindo com suas obrigações atuais (ou seja, solvente em termos de fluxo), os credores tem pouco poder para intervir na situação, já que nenhuma obrigação deixou de ser cumprida.

2.2. Reestruturação de Empresas

Segundo Wruck, (1990, p.425) a dificuldade financeira é resolvida em um ambiente de informações imperfeitas e conflitos de interesse. Os meios de se chegar à solução desse tipo de problema são diversos, e variam ainda conforme a legislação de cada país. Nos Estados Unidos, por exemplo, as reestruturações podem ser feitas por meio de acordos privados fora do ambiente jurídico (*private workouts*), ou podem ser feitos de maneira formal através do Capítulo 11 (Capítulo 11) do *Bankruptcy Reform Act of 1978*. No Brasil, é a Lei de Falências e Concordatas nº 7661 (1945) que regulamenta os pedidos de concordata.

Segundo o Artigo 139 do Título X da Lei de Falências e Concordatas (1945), a concordata no Brasil pode ser classificada como preventiva ou suspensiva, “conforme for pedida em juízo antes ou depois da declaração de falência.” A legislação do Brasil para tratar do assunto de concordatas e falências é muito antiga, e atualmente passa por uma reformulação. Ainda está tramitando no Congresso Nacional uma proposta de mudança da Lei de Falências e Concordatas, que Kayo & Fama (1996, pp.4-6) apresentam resumidamente em seu artigo e que está reproduzida no quadro 1 a seguir.

Quadro 1: *Procedimentos de Concordata no Brasil (Atual e Proposto)*. Fonte: Kayo e Famá, 1996.

Função	Decreto-Lei nº 7661 de 1945	Alterações em Estudo
1. Iniciação dos Procedimentos	<ul style="list-style-type: none"> • Somente pelo Devedor. 	<ul style="list-style-type: none"> • Sem alteração.
2. Custódia da Propriedade	<ul style="list-style-type: none"> • O Juiz faz a nomeação de um comissário dentre os credores quirografários. • O comissário atua somente como fiscal. 	<ul style="list-style-type: none"> • Criação de um comitê (vide item seguinte).
3. Proteção ao Credor	<ul style="list-style-type: none"> • Não existe nenhum comitê. 	<ul style="list-style-type: none"> • “Haveria um Comitê de Recuperação composto por representantes do devedor, dos empregados, dos credores e de técnicos requisitados para implementar o programa de recuperação ou fiscalizar os atos dos administrador judicial na liquidação”. (Háfiez, 1996 apud Kayo e Famá, 1996).
4. Plano de Reorganização (Proposta)	<ul style="list-style-type: none"> • Não existe. 	<ul style="list-style-type: none"> • Apresentação de um plano de recuperação dos devedores ou credores que poderiam vir a ser: <ul style="list-style-type: none"> - Transformação, incorporação, fusão ou cisão da empresa; - Alteração ou substituição do bloco de controle; - Substituição total ou parcial dos administradores; - Aumento do capital social; - Arrendamento; - Constituição de garantias reais e pessoais e - Prazos antecipados e condições especiais de pagamentos para credores que se obriguem a continuar operando normalmente com a empresa, durante toda a recuperação.
5. Revisão da Corte (Juiz)	<ul style="list-style-type: none"> • Determina o processamento, nomeando o comissário ou decreta a falência se não atendidos os pressupostos legais. 	<ul style="list-style-type: none"> • Sem alteração.
6. Plano de Reorganização	<ul style="list-style-type: none"> • Planos de pagamentos: <ul style="list-style-type: none"> - 50% da dívida, se for à vista; - 60%, se for em 6 meses; - 75%, se for em 12 meses; - 90%, se for em 18 meses e - 100%, se for a 24 meses. 	<ul style="list-style-type: none"> • As condições seriam alteradas como seguem: <ul style="list-style-type: none"> - 60% da dívida, se for a 1 ano; - 80%, se for a 2 anos e - 100%, se for a 3 anos.
7. Aprovação	<ul style="list-style-type: none"> • Não precisa ser aprovado por credores. 	<ul style="list-style-type: none"> • Sem alteração.

Para lidar com dificuldades financeiras, as empresas agem de formas variadas de acordo com os interesses das partes envolvidas (acionistas e credores). Segundo Ross, Westerfield & Jaffe (1995, p.621), para resolver tais problemas, as empresas podem passar por uma reestruturação de seus ativos ou por uma reestruturação financeira, conforme exemplos a seguir:

Reestruturação dos ativos:

- Vendendo ativos importantes;
- Combinando-se com outras empresas;
- Reduzindo investimentos e gastos com pesquisa e desenvolvimento.

Reestruturação financeira:

- Emitindo novos títulos;
- Negociando com bancos e outros credores;
- Trocando títulos de dívidas por ações;
- Entrando com pedido de concordata.

As empresas que reestruturam seus ativos procuram focar-se no negócio, às vezes livrando-se de empreendimentos que não são parte da competência essencial da empresa. Outras vezes, as empresas passam por processos de fusão, procurando tornarem-se mais fortes através de alianças com outras empresas. As que escolhem reestruturar seus passivos, podem fazê-lo em geral por meio de negociações privadas ou pedindo concordata formal. Segundo Gilson (1991, p.645), a principal diferença entre os dois processos é que na concordata, todo o processo é supervisionado por uma corte.

Uma outra forma de reestruturação financeira que tem conquistado expressividade nos últimos anos são as concordatas pré-montadas (*pre-packaged bankruptcy*). Segundo Tashjian, Lease & McConnell (1996, p.135), concordatas pré-montadas são uma forma híbrida de reestruturação corporativa combinando algumas das características das reestruturações fora da corte e reestruturação pelo Capítulo 11. Como nas reestruturações privadas, as concordatas pré-montadas têm como característica a negociação da dívida entre os credores e a firma devedora fora do ambiente judicial. Em seguida, como em um pedido de concordata tradicional, a firma entra com o pedido sob os termos do Capítulo 11 para ter uma ratificação da corte.

Muito se tem discutido sobre qual a melhor maneira de se submeter uma firma a uma reestruturação financeira. Segundo Gilson (1991, p.645), como entrar em acordo de maneira privada é apreciavelmente mais barato, uma decisão nesse sentido deveria ser tomada, mas em casos em que um acordo entre a parte credora e a parte devedora se torna impossível, a solução de se pedir concordata passa a ser a única viável.

Quadro 2: Diferença entre as atitudes tomadas por empresas que passaram por dificuldade financeira na década de 1980 e no início da década de 1990, segundo estudo de Gilson (1991).

100%, empresas que descumpriram obrigações na década de 1980	84%, <i>workout</i>	100%, empresas que descumpriram obrigações no início da década de 1990	30%, <i>workout</i>
	16%, Capítulo 11		70%, Capítulo 11

Gilson (1991, p.646) afirma, através de estudos empíricos, que na década de 1980 a reestruturação financeira privada era o método preferido entre as empresas

que passavam por dificuldades financeiras. Segundo dados de sua pesquisa, de 169 empresas que incorreram em não pagamento de dívidas naquela década, 80 foram bem sucedidas em uma negociação privada. Das 89 restantes, pelo menos 62 tentaram uma reestruturação privada antes de pedir concordata².

No início da década de 1990 esses números mudaram de figura. Em uma outra pesquisa de Gilson (1991, p.652), constatou-se que quase 70% das firmas que entraram em dificuldades financeiras entraram com pedido de concordata sem fazer nenhuma tentativa de se reestruturar particularmente com os credores. E para as empresas que tentaram uma negociação privada antes da concordata, apenas três meses se passaram até o pedido de concordata. Na década de 1980 esse número era em média de oito meses.

Não se sabe até que ponto isso não se deva ao advento das concordatas pré-montadas, que sem dúvida passaram a ter destaque entre as formas de reestruturação. O primeiro acordo desse tipo que se tem notícia em uma firma importante foi em 1986, com a *Crystal Oil*. O que surpreendeu nesse caso foi que em poucos meses a *Crystal Oil* conseguiu sair da concordata, evitando os longos períodos que em geral ocorrem em concordatas.

O que em muito influencia a decisão de se fazer uma reestruturação privada ou um pedido formal de concordata nos termos do Capítulo 11 são os custos e benefícios envolvidos em cada uma das duas formas de reestruturação. A seção seguinte tratará desses aspectos com mais detalhes.

² Para obter esses dados, o autor buscou no jornal *The Wall Street Journal* por reportagens que indicassem que a firma passou por alguma tentativa de negociação privada antes de desistir e entrar com o pedido de concordata.

2.3. Custos e Benefícios da Dificuldade Financeira

É de fundamental importância para a gestão de uma empresa saber antecipadamente se sua firma está prestes a passar por uma situação de dificuldades financeiras ou mesmo se seus atuais indicadores econômicos e financeiros apontam para um quadro dessa natureza. O conhecimento antecipado se transforma em um instrumento valioso visando evitar situações desagradáveis de maneira inesperada, como seria uma situação de insolvência, permitindo ao gestor aplicar um plano de recuperação que diminua os custos de um processo deste tipo.

Na escolha entre o pedido formal de concordata ou a tentativa de uma negociação privada, a firma leva em consideração aspectos importantes como o custo e os benefícios envolvidos em cada uma dessas alternativas. Estudos como o de Tashjian, Lease & McConnell (1996, p.155) tentam acrescentar à discussão as questões referentes à concordata pré-montada. Segundo a pesquisa desses autores, algumas medidas utilizadas (como o tempo total gasto na reestruturação e os custos diretos como um percentual do total de ativos antes da dificuldade financeira) apontam para uma opinião de que esses modelos são consideravelmente melhores que a concordata tradicional, mas menos vantajosos que a negociação privada.

Encontra-se na literatura diversos artigos tratando dos custos envolvidos no processo de dificuldade financeira das empresas, mas os benefícios advindos dessa situação só mais recentemente têm sido discutidos (WRUCK, 1990, pp.430-36; GILSON, 1991, pp.649-50). Na questão do endividamento de empresas, Brealey & Myers (1992, pp.437-45) destacam sobremaneira os custos da insolvência financeira.

Na visão dos autores, “a insolvência financeira é onerosa”, e o valor da empresa alavancada é dado por³:

$$VF = VCP + VPBF + VPCIF, \text{ onde:}$$

VF: Valor da firma;

VCP: Valor se totalmente financiada por capitais próprios;

VPBF: Valor presente dos benefícios fiscais da dívida e

VPCIF: Valor presente dos custos de insolvência financeira⁴.

Wruck (1990, p.430) classifica a fórmula apresentada por Brealey & Myers como simples e incompleta, uma vez que ignora a possibilidade da dificuldade financeira resultar em algum benefício, e assim subestima a quantidade de endividamento da firma.

2.3.1 *Benefícios da dificuldade financeira*

Uma das mais importantes conseqüências da dificuldade financeira dentro das empresas é a mudança do sistema de governança corporativa existente. Como afirmam Shleifer & Vishny (1997, p.737), a governança corporativa lida com as maneiras nas quais os fornecedores de capital se asseguram de ter o retorno de seus investimentos, ou, segundo Zingales (1997, p.4), moldam a barganha *ex-post* sobre os excedentes gerados pela firma através de um complexo conjunto de restrições.

³ Essa fórmula visa ajudar a empresa a calcular o seu nível de endividamento ótimo, balanceando os benefícios fiscais do endividamento contra um aumento na probabilidade de ocorrer uma situação de dificuldade financeira.

⁴ VP dos custos de insolvência financeira indica a probabilidade de ocorrência de dificuldades financeiras multiplicado pelos custos esperados dessa situação (diretos e indiretos).

Como afirma Wruck, (1990, p.433) firmas com fraco sistema de governança corporativa são levadas a ter dificuldade financeira. Para se protegerem, os credores dispõem de diferentes formas legais contra expropriações do seu capital. Os credores podem exigir os ativos dados em garantia, entrar com pedido de liquidação da companhia em caso de não pagamento das parcelas da dívida, direito de voto nos casos de reestruturação e ainda o direito de remover maus administradores (SHLEIFER & VISHNY 1997, p.752).

As mudanças na alta administração das firmas em dificuldade financeira são mais freqüentes que nas demais empresas. Segundo Wruck (1990, p.433), as empresas que não estão em dificuldade financeira tem uma rotação anual de altos administradores de cerca de 19%, enquanto que as empresas em dificuldade financeira têm uma rotação anual de 52%.

Em estudo realizado com os altos administradores (*CEO, chairman* e presidente) de 126 empresas listadas na *NYSE (New York and American Stock Exchange)* que descumpriram obrigações de endividamento durante a década de 1980, Gilson (1991, p.651) encontrou que, a despeito da forma de reestruturação adotada (Capítulo 11 ou negociação privada), a rotatividade de membros da alta administração é alta, conforme se pode verificar no Gráfico 1 a seguir. Das empresas que entraram com pedido de reestruturação sob os termos do Capítulo 11, apenas 30% dos administradores originais (computados a partir de um período de dois anos antes da sinalização da dificuldade financeira) permaneceram no cargo. Para as empresas que se reestruturaram de forma privada, a rotatividade foi de 40%. A amostra consistia de 196 administradores originalmente empregados em 126 empresas

listadas na *NYSE* que descumpriram o pagamento de suas obrigações de endividamento no período de 1979-1984. Da amostra, 69 firmas entraram com pedido sob as normas do Capítulo 11 e 57 reestruturaram-se de maneira privada.

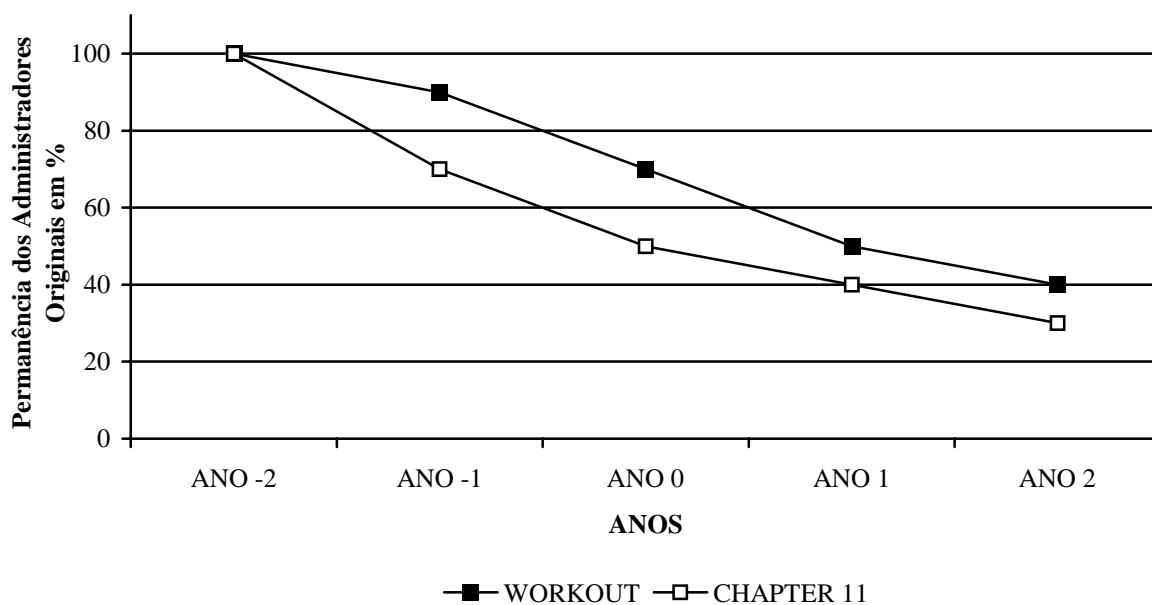


Gráfico 1: Percentual da alta administração original que permanece na firma após um período de dificuldade financeira conforme o métodos de reestruturação financeira. Fonte: Gilson (1991, p.651).

Uma outra consequência direta advinda da dificuldade financeira é a mudança na estrutura e estratégia organizacional. Segundo Wruck (1990, p.434), algumas firmas em dificuldade financeira se submetem a mudanças organizacionais dramáticas como parte de sua recuperação, reajustando o foco estratégico e reestruturando-se.

Apesar de acreditar que no caso geral a reestruturação privada é mais vantajosa que o pedido de concordata dentro do Capítulo 11, Gilson (1991, p.648) enumera quatro vantagens de se pedir reestruturação dentro do Capítulo 11:

- O *Bankruptcy Code* permite às empresas fazer novos endividamentos, que são classificados prioritariamente em comparação a todos os débitos

anteriores. Isso é vital principalmente para empresas que precisam de um alto grau de endividamento para manter o negócio funcionando;

- Os juros sobre as dívidas não seguradas para de aumentar quando a firma está se reestruturando sob os termos do Capítulo 11. Isso permite uma grande economia de recursos importantes em uma hora como a da dificuldade financeira;
- O *Bankruptcy Code* protege as empresas que estão se reestruturando do forte assédio dos credores, permitindo que a empresa funcione com menos turbulência;
- É mais fácil obter aprovação do plano de reestruturação dentro do Capítulo 11 porque as regras de votação são menos restritivas: a aprovação se dá por maioria (50% em quantidade, desde que represente dois terços em valor) dos *claimholders*.

2.3.2 *Custos da dificuldade financeira*

Comparativamente, como já foi mencionado a partir do estudo de Gilson (1991), os custos de reestruturação através de *workouts* são significativamente menores que através de pedido formal de concordata sob os termos do Capítulo 11, apesar das vantagens mencionadas na seção anterior. No entanto, os custos de cada uma dessas formas de reestruturação devem ser analisados cuidadosamente antes de se decidir por uma ou outra.

Os custos de reestruturação são classificados em custos diretos, também chamados de *out-of-pocket*, e custos indiretos. Os custos diretos são sempre facilmente mensurados, enquanto que os custos indiretos, embora conhecidos, são difíceis de quantificar.

Entre os custos diretos encontram-se as taxas legais e honorários dos profissionais envolvidos, como advogados, consultores e auditores, que são geralmente calculados em horas. Como os processos de reestruturação costumam demorar longos meses⁵, esse custo tende a ser um dos mais elevados. Uma solução potencial é apresentada por Gilson (1991, p.647), que sugere a distribuição de ações para esses profissionais da mesma forma que é feita entre credores e acionistas, incentivando-os a interessarem-se em manter o valor da firma o maior possível.

Entre os custos indiretos da dificuldade financeira, destacam-se principalmente os custos advindos de oportunidades de negócio perdidas. Como destaca Wruck (1990, p.439), uma firma que se encontra em dificuldade financeira é atingida de três diferentes formas:

- (i) Ela perde o direito de tomar certas decisões sem uma aprovação legal. É o caso das empresas que estão se reestruturando sob os termos do Capítulo 11, que devem obter aprovação da corte que supervisiona a falência caso queiram tomar decisões como vender ativos;

⁵ Segundo Gilson, John & Lang, “Troubled Debt Restructurings: an empirical study of private reorganizations of firms in default”, *Journal of Financial Economics* 26 (1990), o tempo médio gasto com reestruturação no Capítulo 11 foi de mais de 20 meses, enquanto o tempo médio para reestruturação privada foi de cerca de 15 meses.

- (ii) A dificuldade financeira lança no mercado uma má notícia que eventualmente prejudica as vendas da empresa envolvida, e conseqüentemente pode aumentar os custos de produção, e
- (iii) A alta administração da empresa tem grande parte de seu tempo sendo consumido na solução dos problemas financeiros, o que tira seu foco para oportunidades que podem alavancar o valor da firma.

Gilson (1991, pp.647-8) afirma que um grande desperdício de ativos é possível quando uma firma entra com pedido nos termos do Capítulo 11. Isso porque o *Bankruptcy Code* requer que um juiz aprove as decisões financeiras mais importantes da empresa nesse período, e há uma influência muito grande da corte em como os negócios internos da companhia são conduzidos.

Além disso, vários procedimentos obrigatórios pelo *Bankruptcy Code* podem atrasar tomadas de decisão dentro da empresa. Situações em que se deseja por exemplo tomar novos empréstimos, ou vender ativos devem ser encaminhadas à corte e aos credores para aprovação, e isso pode causar alguns meses de demora.

Devido a todas essas questões envolvendo benefícios e custos para empresas e credores em cada uma das formas de reestruturação é que conseguir prever antecipadamente tais situações pode evitar uma série de problemas. Na seção seguinte serão apresentados alguns dos modelos de previsão de falência mais importantes já desenvolvidos .

2.4. Modelos de Previsão de Falência

Desde o primeiro trabalho que se tem notícia sobre a previsão de falência em empresas realizado em 1932 por Fitzpatrick, muito se avançou em estudos dessa natureza. Apesar de não termos ainda hoje uma teoria definitiva sobre o assunto, diversos estudos empíricos com empresas de vários países do mundo têm sido realizados sempre com o intuito de buscar conhecer antecipadamente quais os determinantes de uma empresa em vias de entrar em processo de insolvência ou mesmo simplesmente se determinada empresa corre ou não o risco de entrar em um processo como este. A motivação para esses estudos vem da necessidade de se tomar atitudes que busquem evitar uma possível liquidação de empresas ou ainda para resguardar os interesses de quaisquer que sejam os envolvidos com elas, que pretendem não ter prejuízos com eventuais rumos desastrosos que possam estar seguindo.

Mas foi apenas cerca de 3 décadas depois de Fitzpatrick que ferramentas estatísticas passaram a fazer parte dos estudos de previsão de insolvência de empresas. Esses modelos foram acentuadamente mais eficientes em sinalizar com antecedência a situação indesejável, permitindo que ações corretivas fossem tomadas para se evitar a insolvência.

Inicialmente, a técnica estatística adotada para realizar as previsões de falência de empresas foi a Análise Discriminante Univariada (ADU), através do estudo de Beaver (1967). Em seguida, Altman (1968) introduziu a Análise Discriminante Linear (ADL). Segundo Altman (1983, p.102), a escolha se deu após cuidadosamente considerar a natureza do problema e o propósito de sua análise.

Segundo Hair e outros autores (1998, p.14), a ADL pode ser usada quando a variável dependente é não-métrica e as variáveis independentes são métricas. Essa análise é útil em situações em que a amostra pode ser dividida em diversos grupos distintos baseados na característica da variável dependente. O maior objetivo dessa técnica é entender as diferenças existentes entre esses grupos, e prever a probabilidade de que uma entidade pertença a um grupo ou aos demais. No caso dos estudos de predição de falência, a variável dependente é dicotômica (solvente ou insolvente) e não-métrica, atendendo às necessidades do modelo.

A partir dos dois primeiros modelos apresentados por Altman (1968 e 1979), o *Modelo Z-Score* e o *Modelo Zeta-Analysis*, diversos outros estudos sobre modelos de previsão de falência foram publicados. Nesta seção serão apresentados alguns dos principais modelos utilizados ao longo desses anos de acordo com a técnica empregada. Além desses dois modelos de Altman, serão apresentados modelos que utilizam a regressão logística (RL), técnica introduzida por Ohlson em 1980 e as recentes utilizações de Redes Neurais Artificiais (RNA) na análise de insolvência.

2.4.1 *Estudos Univariados*

2.4.1.1 *Estudo de Beaver (1967)*

O estudo precursor de Beaver tratava como empresas com problemas aquelas que não cumpriam com suas obrigações financeiras nas datas de pagamentos. Ele estudou empresas entre os anos de 1954 e 1964, cujos ativos variavam entre \$0,6 milhão de dólares e \$45 milhões de dólares. Ele buscou controlar seu estudo usando

uma amostra pareada, respeitando aspectos como indústria de atuação e porte da firma. Beaver justificava essa preocupação com o argumento de que entre diferentes indústrias, o mesmo valor numérico para uma variável pode representar diferentes probabilidades de dificuldades, e dados indicadores iguais, firmas menores podem ter uma maior probabilidade de falência.

Beaver adotou 30 indicadores financeiros, e para cada empresa computou-os para os últimos cinco anos disponíveis. De todos estes índices analisados, seis foram selecionados como sendo aqueles que contém uma maior quantidade de conteúdo informacional a respeito da situação de dificuldade financeira (ou não) da empresa.

Os índices relacionados por Beaver foram os seguintes:

- Fluxo de Caixa/Total de Endividamento;
- Lucro Líquido/Ativo Total;
- Passivo Circulante + Exigível de Longo Prazo/Ativo Total;
- Capital de Giro/Ativo Total;
- Índice de Liquidez Corrente e
- Intervalo sem Crédito.

Beaver usou 2 técnicas de classificação na realização de seu experimento: a primeira foi comparação de médias, e a segunda foi uma espécie de classificação dicotômica. Com a comparação de médias, Beaver conseguiu verificar a deterioração dos índices à medida que se tomavam dados financeiros cada vez mais próximos da data da dificuldade financeira no caso das empresas insolventes. Para as empresas solventes, os dados se mostravam relativamente constantes ao longo do tempo.

O teste de classificação foi realizado da seguinte forma: Beaver ordenou cada uma das 30 variáveis do estudo para cada um dos grupos e determinou visualmente qual deveria ser o ponto de corte que resultaria em um menor número de

classificações incorretas. Assim, o percentual de classificações incorretas foi tomado como o índice de medida da capacidade de previsão de cada variável. Desta forma, é possível observar que o estudo de Beaver considera que o custo de classificar incorretamente uma empresa insolvente como solvente (erro do Tipo I) é o mesmo que classificar incorretamente uma empresa solvente como insolvente (erro do Tipo II). Isto é uma inverdade, pois o erro do Tipo I é na maioria das vezes mais caro do que o erro do Tipo II. A figura 3 ilustra como era a metodologia de escolha do ponto de corte realizada por Beaver em seu estudo.

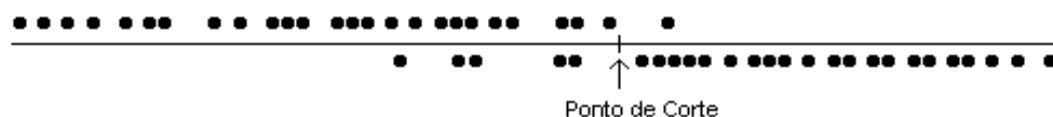


Figura 3: Exemplo da classificação dicotômica feita por Beaver (1967).

Assim, analisando os índices um a um, Beaver chegou à conclusão de que o melhor indicador financeiro para discriminar empresas solventes de empresas insolventes era o fluxo de caixa/total de endividamento. Essa conclusão poderia ser diferente caso ele tivesse levado em consideração as diferenças entre os prejuízos causados por prever incorretamente uma empresa solvente como insolvente e vice-versa.

Beaver, nas conclusões finais a respeito de sua pesquisa alega que a maior contribuição alcançada pelo estudo foi a de fornecer uma metodologia para a análise de dados financeiros de empresas, e que é possível prever a ocorrência de situações de insolvência com até cinco anos de antecedência. Ele salienta que alguns são melhores

para prever a situação de solvência do que a situação de insolvência, e que nem todos são capazes de prever insolvência com a mesma eficiência.

Um ano mais tarde, Beaver (1968) publicou um outro estudo usando técnica univariada de análise de dados. No estudo, ele investiga até que ponto as mudanças nos preços de mercado das ações de empresas podem ser usadas para prever uma situação de falência. Foi utilizada uma variável para medir a mudança de preço de mercado de uma ação⁶ e seu comportamento foi comparado ao comportamento de alguns indicadores que haviam se mostrado eficientes em prever falência. Beaver encontrou que havia uma certa correlação entre a variável de mercado e as variáveis contábeis (o que sugere a utilidade de uma nova medida que leve em consideração a avaliação de mercado da empresa), mas que mesmo assim a capacidade discriminatória desta nova variável não era tão boa quanto às do estudo de 1967.

2.4.1.2 *Estudo de Kanitz (1976)*

A fim de responder a perguntas do tipo “é possível prever a falência de uma empresa?”, ou “quais são os melhores indicadores de previsão?”, ou “os balanços das empresas brasileiras são fidedignos?”, Kanitz (1976) testou para empresas brasileiros a eficiência de análise de índices financeiros. A quantidade de indicadores selecionados para o estudo foi excessiva: Kanitz listou 516 indicadores, todos sugeridos pela

⁶ No estudo de Beaver (1968), a variável que mede a mudança de preço de mercado de uma ação é R_{it} , que é igual a $\frac{D_{it} + P_{it} - P'_{it-1}}{P'_{it-1}}$, onde P_{it} é o preço da ação i no tempo t , D_{it} é o pagamento de dividendos da ação i entre o tempo $t-1$ e t , e P'_{it-1} é o preço da ação i no tempo $t-1$, ajustado para mudanças de capital (exemplo: desdobramento de ações e pagamento de dividendos).

literatura especializada consultada quando da elaboração do estudo⁷, sem descartar nenhum dos indicadores encontrados, mesmo que seu significado prático pudesse parecer pouco contributivo.

Apesar das críticas a que podem estar sujeitos os demonstrativos financeiros das empresas (muitas delas justamente), Kanitz buscou acima de tudo provar que um balanço quando bem analisado pode fornecer indicações valiosas a respeito da saúde financeiras das empresas. Kanitz analisou os balanços das empresas buscando obter informações a partir da posição relativa da empresa, projetando balanços para períodos futuros e elaborando alguns indicadores com base nos demonstrativos financeiros.

Para montar as posições relativas dos indicadores das empresas estudadas, Kanitz fez uso do conceito de centil, ou seja, para cada empresa foi determinado o percentual de outras empresas que seriam superadas para cada índice analisado. Essa forma simplista de verificar a probabilidade de que uma empresa venha a entrar em falência se baseia no seguinte argumento: uma empresa que está, por exemplo, no decil 15 (acima de 15% das empresas da amostra) deverá entrar em falência, mas não antes das demais que estão abaixo dela. Como não há essa lógica no processo de falência de empresas, pode-se considerar esse método muito frágil.

Para se saber quais índices são os melhores para a previsão de falência, Kanitz tomou aqueles que listavam como piores da lista aquelas empresas que efetivamente haviam quebrado. Se não houvesse uma separação clara quanto à situação das empresas, aquele indicador era tomado como inútil para prever falência.

⁷ Ver referências na Página 8 de Kanitz (1976).

Para realizar a análise através dos balanços projetados das empresas, Kanitz pressupôs como critério o de que as empresas tomariam nos dois anos seguintes as mesmas decisões tomadas no último ano de dados disponíveis. Caso a repetição das decisões tomadas não deteriorasse os índices das empresas nos anos seguintes, poder-se-ia inferir que as decisões da diretoria estavam sendo satisfatórias e que a empresa não sofreria risco de falência. Do contrário, caso a repetição das decisões trouxesse quedas na qualidade dos indicadores, aquela empresa representava uma iminente candidata à falência.

Na seleção da amostra, Kanitz usou o critério de emparelhamento de empresas, tomando 21 empresas solventes e 21 insolventes, usando o porte e a indústria de atuação da empresa como variáveis de controle. O período utilizado no estudo compreendeu os anos entre 1972 e 1974, sendo analisados os balanços dos dois anos anteriores à data da falência. A análise estatística utilizada (diferença de médias), aponta para diversos índices que individualmente servem como discriminadores de empresas em boa situação financeira de empresas que foram à falência. O surpreendente número de 81 indicadores se mostrou significativo ao nível de 5% na discriminação entre os dois grupos de empresas.

Apesar da técnica utilizada em seu estudo ser atualmente considerada primitiva para um estudo desta natureza, a pesquisa de Kanitz deve ser tratada como um marco importante na pesquisa brasileira de previsão de falência de empresas. Na época em que ele escreveu seu trabalho, publicado como Tese de Livre Docência para a Faculdade de Economia e Administração da Universidade de São Paulo (FEA/USP), Altman já havia realizado seu modelo usando a análise discriminante

multivariada. Um estudo com dados brasileiros usando essa técnica então moderna só seria realizado três anos mais tarde, pelo próprio Altman em conjunto com dois professores brasileiros⁸.

2.4.2 *Modelos Discriminantes Multivariados*

2.4.2.1 *Modelo Z-Score de Altman (1968)*

O modelo *Z-Score* proposto por Altman em 1968 foi obtido através de um estudo que tinha como propósito abordar a qualidade das técnicas analíticas que usavam indicadores contábeis para medir o desempenho de um negócio. Ele então fez uso da previsão de falência corporativa como um instrumento para comprovar sua tese de que tais indicadores, quando analisados sob uma perspectiva estatística multivariada, oferecem uma melhor significância estatística do que as então comuns técnicas de comparação de indicadores com padrões.

Antes da utilização de técnicas quantitativas, os estudos realizados por pesquisadores ou por agências de classificação de risco, utilizavam sobremaneira dados e análise meramente qualitativas. Altman partiu de alguns estudos pioneiros, entre eles o de Beaver (1967) para concluir que indicadores contábeis possuem bom potencial para serem usados como variáveis de previsão de falência de empresas. Ele observou também que índices que retratam lucratividade, liquidez e solvência são prevalentes nos estudos como aqueles mais significativos, embora uma ordem de

⁸ Este estudo será discutido na seção seguinte.

prevalência fosse difícil de ser determinada uma vez que em cada estudo havia um índice diferente como sendo aquele que fornecia indícios mais fortes a respeito da situação da empresa.

Altman questionava a capacidade dos estudos anteriores para determinar o potencial de empresas estarem encaminhando-se para a falência por que a metodologia empregada por seus predecessores era essencialmente univariada. Na interpretação de Altman, análise univariada de indicadores é muito “susceptível a interpretações erradas e é potencialmente causadora de confusão” (Altman, 1968, p.591). Ele cita como exemplo uma empresa que possui baixa lucratividade ou antecedentes de solvência, que em estudos univariados não seria classificada como saudável, mas que possui uma excelente liquidez, que a coloca em uma situação não grave. Segundo seu argumento, só uma análise multivariada poderia combinar diversas medidas em um modelo de previsão que fizesse sentido.

Altman usou ainda algumas das conclusões a que Beaver havia chegado em seu estudo um ano antes, como a de que dados contábeis podem prever falência com até cinco anos de antecedência. Segundo Altman (1983, p.99), na maioria das vezes, os problemas fundamentais que levam o negócio a situações de dificuldade financeira estão dentro da própria empresa. A questão então passara a ser quais indicadores seriam mais importantes para detectar o potencial de uma empresa estar caminhando em direção a uma situação de dificuldade financeira, e quais pesos atribuir para cada um desses indicadores.

A ADL era ainda muito recente no campo das Finanças, onde predominava a Análise de Regressão Múltipla (ARM). A ADL no entanto possuía grande aplicação

nas ciências biológicas e nas ciências comportamentais⁹. Em Finanças, esta técnica havia sido aplicada em estudos de avaliação de crédito ao consumidor (DURANT, 1941 e MYERS & FORGY, 1963) e em classificação de investimentos (WALTER, 1959).

A ADL busca classificar corretamente uma observação em diversos grupos (2 ou mais) de acordo com as características individuais de cada observação. A técnica procura essencialmente classificar ou fazer previsão em situações onde a variável dependente é qualitativa (masculino/feminino, solteiro/não solteiro, saudável/doente, etc.). Portanto, estabelecer claramente os critérios de classificação entre os grupos existentes é fundamental. Após o estabelecimento dos grupos coletam-se dados para os sujeitos de cada um dos grupos. A técnica de ADL procura então extrair dos dados uma combinação linear das características dos sujeitos (variáveis independentes) que melhor discriminem os grupos. Se um sujeito em particular (no caso, uma empresa) possui características (indicadores financeiros) que possam ser quantificados para todas as companhias sob análise, a técnica da ADL determina um conjunto de coeficientes que atribuídos a cada variável independente possui poder de discriminação. Quando então esses coeficientes são aplicados em um conjunto de indicadores de uma empresa, obtém-se uma base para classificação daquela empresa em um dos grupos mutuamente exclusivos definidos *a priori*.

De acordo com Altman, a ADL possui uma importante vantagem em comparação à ADU: a primeira considera um conjunto de características comuns às empresas sob estudo bem como as interações dessas propriedades, enquanto que a

⁹ Para conhecer mais sobre os usos e aplicações da ADL, ver Cochran (1964).

técnica univariada só pode considerar os indicadores usados para o processo de classificação dentro dos grupos um de cada vez.

A figura a seguir, extraída de Altman (1983, p.104), mostra uma análise discriminante onde duas variáveis (lucratividade e liquidez) são plotadas em um gráfico para empresas saudáveis (apontadas com um x) e empresas não saudáveis (apontadas com um o). A técnica usada nos modelos discriminantes consiste em atribuir pesos que separarão tanto quanto possível as médias dos valores para cada grupo e ao mesmo tempo tentará minimizar a distância estatística de cada observação (cada x e cada o) e a média do próprio grupo. A linha AB da figura é a que melhor discrimina os dois grupos.

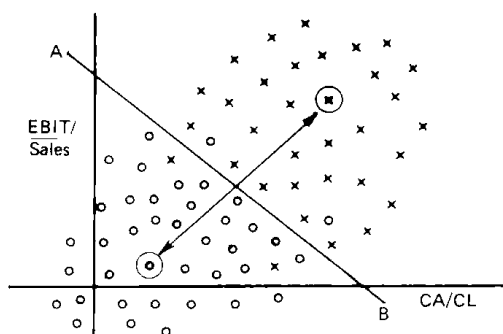


Gráfico 2: Análise Discriminante Linear. Exemplo extraído de Altman (1983, p.104). No gráfico, as empresas saudáveis são denotadas com um x , e as empresas não saudáveis com um o . As médias dos grupos estão circuladas, e a linha AB é a que melhor discrimina os dois grupos.

A função discriminante obtida é da forma $Z = k_1X_1 + k_2X_2 + \dots + k_nX_n$, onde k_1, k_2, \dots, k_n são os coeficientes de discriminação e X_1, X_2, \dots, X_n são as variáveis independentes. Esta função transforma os valores de cada variável em um escore discriminante Z , que então é usado para classificar o objeto de estudo.

Dadas as fortes correlações existentes entre as variáveis medidas nas empresas que passam por situação semelhante, Altman esperava que seu modelo final possuísse poucas variáveis, mantendo ainda um elevado grau de informação e um alto poder de discriminação entre os grupos. Essa característica, chamada colinearidade, ao mesmo tempo em que permite que um número menor de variáveis deva ser mantida no modelo fornecendo a mesma quantidade de informação é também uma causadora potencial de problemas em trabalhos dessa natureza. Segundo Neter *et alii* (1996), os coeficientes estimados no modelo de regressão tendem a ter grande variabilidade amostral quando as variáveis independentes são altamente correlacionadas, causando uma baixa qualidade de informação acerca dos coeficientes da regressão verdadeira.

Para desenvolver seu modelo discriminante, Altman tomou 66 empresas, sendo 33 para cada um dos grupos de interesse. O grupo de empresas insolventes era formado por empresas que haviam pedido concordata sob os termos do Capítulo 11 durante os anos de 1946 a 1965. Por conta de limitações na disponibilidade de dados, Altman se viu obrigado a utilizar um intervalo de tempo tão grande. Ele mesmo afirma que a escolha por um período de 22 anos não é o melhor procedimento, uma vez que os índices médios mudam ao longo do tempo.

Altman não fez uso de um grupo de empresas insolventes completamente homogêneo: o valor dos ativos das empresas estudadas variava entre \$ 0,7 milhão de dólares a \$ 25,9 milhões de dólares. Para compor o grupo de empresas solventes, ele tomou uma amostra pareada, buscando controlar efeitos como indústria de atuação e porte. Este grupo de empresas tinha ativos que variavam entre \$ 1 milhão e \$ 25 milhões de dólares, escolhidas principalmente para excluir da amostra empresas muito

pequenas e muito grandes e acompanhar o intervalo disponível das empresas do primeiro grupo.

A fim de obter seu modelo, Altman selecionou um grupo de 22 variáveis potencialmente explicativas do fenômeno insolvência. Essas variáveis eram classificadas em 5 grupos distintos, incluindo índices de liquidez, lucratividade, alavancagem, solvência e atividade. Elas foram escolhidas sobretudo por sua popularidade dentro da literatura disponível sobre o assunto, potencial relevância para o estudo e ainda alguns indicadores “novos” lançados por seu estudo.

Das 22 variáveis selecionadas para o estudo, o modelo resultou em 5 variáveis para compô-lo. A fim de chegar nestas variáveis definitivas, os seguintes procedimentos foram executados:

- Observação da significância estatística de várias funções alternativas, incluindo a determinação da contribuição relativa de cada variável independente;
- Avaliação da inter-correlação entre as variáveis relevantes;
- Observação da precisão de previsão dos diversos perfis, e
- Julgamento do analista.

Altman destaca em seu estudo que as variáveis mais significativas medidas independentemente não fizeram parte de seu modelo final. A função discriminante fornecida pela ADL foi a seguinte:

$$Z = 0,012 X_1 + 0,014 X_2 + 0,033 X_3 + 0,006 X_4 + 0,999 X_5, \text{ onde, segundo}$$

Caouette, Altman & Narayanan (1999, pp.131-3):

X₁, Capital de Giro/Ativo Total:

O índice de capital de giro/ativo total, frequentemente encontrado em estudos de problema empresariais, é uma medida dos ativos líquidos de uma empresa em relação a sua capitalização total. Define-se capital de giro como a diferença entre o ativo circulante e o passivo circulante. Características de liquidez e porte também são levadas em consideração. Geralmente, uma empresa que sofre perdas operacionais consistentes terá diminuição do ativo circulante em relação ao ativo total. Dos três índices de liquidez avaliados, este provou ser o mais valioso.

X₂, Lucros Retidos/Ativo Total:

Lucros retidos são o valor total dos lucros e/ou prejuízos reinvestidos em uma empresa ao longo de sua vida. A conta de lucros retidos está sujeita a manipulação por pseudo-reorganizações societárias e declarações de dividendos. Apesar de tais fatos não serem evidentes neste estudo, é concebível que uma reorganização ou distribuição de dividendos substanciais possam criar uma distorção. Neste caso, as contas devem sofrer ajustes apropriados. A idade de uma empresa está implícita nesta razão: uma empresa relativamente jovem, por exemplo, provavelmente terá LR/AT baixo, porque não teve tempo para acumular lucros retidos. Assim, pode-se argumentar que esta análise faz discriminação contra empresas jovens. Na verdade, contudo, esta distorção é perfeitamente razoável, já que a incidência de quebra é muito mais alta em empresas jovens¹⁰. Essa medida foi uma das citadas por Altman como “novos” indicadores.

¹⁰ Em 1996 aproximadamente 45% das empresas que quebraram, o fizeram nos primeiros 5 anos de sua existência (Dun & Bradstreet, 1997 apud Caouette, Altman & Narayanan, 1999).

X₃, Lucros Antes de Juros e Imposto de Renda/Ativo Total:

A razão entre os lucros antes de juros e imposto de renda e o ativo total (LAJIR/AT) é uma medida da produtividade dos ativos das empresas independentemente de quaisquer fatores tributários ou de alavancagem. Como a existência de uma empresa se baseia, em última análise, no poder de ganho de seus ativos, este índice parece especialmente adequado a estudos que lidem com a quebra de empresas. Ademais, a insolvência ocorre quando o passivo total de uma empresa supera uma justa avaliação de seus ativos, tais como determinada pelo poder de ganho de tais ativos.

X₄, Valor de Mercado do Patrimônio Líquido/Valor Escritural do Passivo Total:

O patrimônio líquido se mede pelo valor combinado de mercado de todas as ações preferenciais e ordinárias, ao passo que o passivo inclui itens circulantes e de longo prazo. O índice entre o valor de mercado das ações e o valor escritural do passivo total mostra o quanto os ativos de uma empresa podem cair de valor (igual ao valor de mercado do patrimônio somado ao endividamento) antes que seu passivo supere seu ativo e a empresa se torne insolvente. Por exemplo, uma empresa com valor de mercado do patrimônio líquido de \$1.000 e endividamento de \$500 poderia sofrer queda de 2/3 do valor de seu ativo ($\$1.000 + \$500 = \$1.500$ de maneira que 2/3 do ativo equivalem a \$500) antes de se tornar insolvente. Contudo, uma empresa com valor de patrimônio líquido de \$250 se tornará insolvente se seu ativo sofrer desvalorização de apenas 1/3. Este índice acrescenta uma dimensão de valor de mercado que outros estudos de quebra deixaram de considerar.

X_5 , *Vendas/Ativo Total*:

O índice do giro do ativo é um índice financeiro comum que ilustra a capacidade de geração de faturamento dos ativos de uma empresa. É uma medida da capacidade dos administradores da empresa para lidar com as condições competitivas. Baseado apenas na medida de sua relevância estatística univariada, não teria sequer aparecido no modelo. Contudo, graças a sua relação exclusiva com as demais variáveis do modelo o índice vendas/ativo total fica em segundo lugar como contribuição à capacidade discriminatória geral do modelo.

A tabela a seguir apresenta as médias das variáveis para os grupos de empresas solventes e insolventes e o teste de significância F para a habilidade discriminante individual de cada variável.

Variável	Média do Grupo de Empresas Insolventes	Média do Grupo de Empresas Solventes	Estatística F
	n = 33	n = 33	
X_1	-6,1%	41,4%	32,60*
X_2	-62,6%	35,5%	58,86*
X_3	-31,8%	15,3%	26,56*
X_4	40,1%	247,7%	33,26*
X_5	150,0%	190,0%	2,84

* Significante ao nível de 0,1%.

Tabela 1: Média das Variáveis e Teste de Significância. Fonte: Altman, 1968.

As variáveis X_1 a X_4 foram todas significantes ao nível de 0,001 (0,1%) indicando diferenças significativas entre os grupos. A variável X_5 foi incluída no

modelo por conta de sua alta contribuição individual ao poder de discriminação total da função¹¹ Z.

Os resultados exploratórios do teste empírico de Altman podem ser resumidos como mostra a tabela a seguir. Esses resultados são do tipo classificatórios, e não do tipo de previsão, pois neste estágio o único interesse é o de explorar se o modelo classifica adequadamente as empresas usadas na geração da equação discriminante.

Grupo Efetivo	Número de Casos	Grupo Previsto	
		1	2
Grupo 1 (insolventes)	33	31 94%	2 6%
Grupo 2 (solventes)	33	1 3%	32 97%

Precisão de Classificação Geral: 95%

Tabela 2: Resultado do estágio classificatório do modelo discriminante de Altman medido um ano antes da falência. Fonte: Altman, 1968.

Grupo Efetivo	Número de Casos	Grupo Previsto	
		1	2
Grupo 1 (insolventes)	33	23 72%	9 28%
Grupo 2 (solventes)	33	2 6%	31 94%

Precisão de Classificação Geral: 82%

Tabela 3: Resultado do estágio classificatório do modelo discriminante de Altman medido dois anos antes da falência. Fonte: Altman, 1968.

As tabelas 2 e 3 apresentadas acima mostram a precisão de classificação do modelo de Altman. Elas são úteis para determinar se na prática o modelo é bom.

¹¹ A variável X₅, das cinco incluídas no modelo, teve o segundo maior grau de contribuição relativa, só perdendo para a variável X₃.

Com os dados das empresas um ano antes delas pedirem falência, o modelo foi capaz de classificar corretamente 95% das empresas, e para os dados de dois anos antes, foi capaz de classificar corretamente 82% delas. O cálculo dessa precisão global é fruto de duas precisões distintas, que Altman chamou de precisão Tipo I e precisão Tipo II. A precisão Tipo I é a precisão do modelo em classificar como insolventes as empresas que quebraram, enquanto que a precisão Tipo II é a precisão do modelo em classificar como solventes as empresas que não quebraram. A precisão geral é a combinação dos dois tipos, ou seja, quantas empresas do total de analisadas foram classificadas corretamente. Conforme afirma Caouette,

a precisão Tipo I é vista como sendo mais importante do que a Tipo II, porque a incapacidade de identificar uma empresa em processo de quebra custará ao aplicador muito mais do que o custo de oportunidade de rejeição de uma empresa saudável como quebra em potencial (1999, p.133).

Outros quatro testes foram ainda realizados por Altman como forma de validar seu modelo discriminante. Isso foi importante, pois utilizar a mesma amostra para construir e validar o modelo não é uma boa idéia. É importante ter uma amostra de validação diferente da amostra usada para construir o modelo. Assim, como primeiro teste de validação Altman utilizou 5 sub-amostras, todas de tamanho 16, da amostra inicial de empresas. Os critérios utilizados por ele para chegar nessas 16 empresas foram:

- Amostragem aleatória;
- Seleção alternada de empresas, começando da primeira empresa;
- Seleção alternada de empresas, começando da segunda empresa;
- Seleção das empresas de 1 a 16; e
- Seleção das empresas de 17 a 32.

Os resultados dos testes com estas cinco sub-amostras está apresentado na tabela abaixo.

Replicação	Percentual de Classificações		Valor de t
	Corretas		
1	91,2		4,8*
2	91,2		4,8*
3	97,0		5,5*
4	97,0		4,5*
5	91,2		4,8*
Média	93,5		5,1*
Número total de observações por replicação.....			34

* Significante ao nível de 0,1%

Tabela 4: Precisão dos testes de validação para as sub-amostras. Fonte: Altman, 1968.

O teste t foi aplicado para testar a significância dos resultados obtidos. A fórmula de t utilizada foi

$$t = \frac{\text{proporção correta} - 0,5}{\sqrt{\frac{0,5(1-0,5)}{n}}}$$

Como segundo teste de validação foi selecionada uma nova amostra de 25 empresas insolventes. Estas empresas possuíam um porte (medido pelo tamanho de seus ativos) semelhante às empresas da amostra original. Usando os parâmetros do modelo discriminante foram encontrados os seguintes resultados, como apresenta a tabela a seguir.

		Preditas	
		Insolventes	Solventes
Insolventes (Real)		24	1
	Número de Predições Corretas	Percentual de Predições Corretas	Percentual de Erros
Tipo I	24	96	4
			N
			25

Tabela 5: Precisão dos testes de validação para nova amostra de empresas insolventes. Fonte: Altman, 1968.

O resultado deste teste surpreendeu Altman, pois ele não esperava encontrar uma amostra que lhe desse uma maior percentual de acertos (96% contra 94%) do que a amostra original usada para a construção do modelo.

O próximo teste realizado para validar o modelo usou uma nova amostra de empresas que apresentavam resultados negativos em seus balanços mas que não haviam pedido concordata sob os termos legais norte-americanos. Ele selecionou 66 empresas que tivessem reportado perdas nos anos de 1958 e 1961. Para esta amostra Altman não fez restrição ao porte da empresa, mantendo apenas a restrição do setor de manufatura. Os resultados da previsão realizada pelo modelo são os seguintes.

		Preditas		
		Insolventes	Solventes	
Solventes (Real)		14	52	
	Número de	Percentual de	Percentual de	N
	Predições Corretas	Predições Corretas	Erros	
Tipo II	52	79	21	66

Tabela 6: Precisão dos testes de validação para nova amostra de empresas solventes, mas que reportaram prejuízos nos anos de 1958 e 1961. Fonte: Altman, 1968.

Para este caso o modelo discriminante classificou corretamente 79% das empresas. Este resultado também impressionou Altman, uma vez que as empresas utilizadas por ele na amostra poderiam ser consideradas como abaixo da média quanto ao critério desempenho.

Finalmente, o quarto teste realizado por Altman buscava verificar a capacidade de seu modelo em prever a situação das empresas no longo prazo. Para isso, tomou os cinco últimos anos de exercícios da amostra original de empresas e buscou verificar

como seu modelo classificava cada uma destas empresas. Os resultados deste último teste estão mostrados na tabela a seguir.

Anos antes da Falência	Acertos	Erros	Percentual de Acertos
1 (n=33)	31	2	95
2 (n=32)	23	9	72
3 (n=29)	14	15	48
4 (n=28)	8	20	29
5 (n=25)	9	16	36

Tabela 7: *Precisão dos testes de classificação para as empresas da amostra original no longo prazo.*
 Fonte: Altman, 1968.

Os resultados acima mostram que o modelo era eficiente em classificar empresas quanto ao seu status de solvente ou insolvente para até dois anos antes da falência, mas que os resultados não eram tão confiáveis assim para anos seguintes.

Os resultados do teste empírico permitiram a Altman estabelecer um ponto de corte: ele percebeu que todas as empresas com *Z-Score* maior que 2,99 caíram no setor de solventes, enquanto que todas as empresas com *Z-Score* menor que 1,81 eram insolventes. Os valores dentro desse intervalo foram chamados por ele de “zona da ignorância” (Altman, 1983, p.120), pois são muito suscetíveis a erros de classificação.

O *Modelo Z-Score* foi muito contributivo com os estudos de previsão de falência. Ele demonstrou que a ADL era mais adequada e robusta em suas conclusões que a ADU, muito embora seus resultados para períodos maiores que dois anos não tenham sido satisfatórios.

2.4.2.2 *Modelo ZETA[®] Analysis (1977)*

O *Modelo ZETA[®]*, desenvolvido por Altman, Haldeman & Narayanan em 1977 sucedeu o anterior *Modelo Z-Score*. Esse novo modelo, desenvolvido conjuntamente com uma empresa financeira privada, tinha capacidade de previsão de 90% para dados um ano antes da falência e de 70% para dados cinco anos antes. Dessa vez, procurou-se dar maior destaque a empresas de maior porte, e foram escolhidas aquelas que tinham uma média dos ativos totais de cerca de \$100 milhões de dólares dois anos antes da insolvência. Para relembrar, o modelo anterior foi feito com empresas que tinham no máximo \$25 milhões de dólares em ativos totais, e neste estudo a menor empresa da amostra possuía ativo total avaliado em \$20 milhões de dólares.

As razões apresentadas por Altman para desenvolver um novo modelo quase dez anos depois da introdução do escore Z foram baseadas na verificação que, após esse período, houve uma mudança de tamanho das empresas que entravam em falência, exigindo que um novo modelo que contemplasse empresas de grande porte fosse desenvolvido. Foram feitas também algumas correções no modelo anterior de modo a comportar uma indústria especialmente vulnerável: a do varejo. Os autores preocuparam-se também em incorporar alguns dos refinamentos que as técnicas estatísticas discriminantes haviam alcançado nos últimos anos, na esperança de que assim pudessem obter resultados mais precisos do que os obtidos anteriormente.

A amostra analisada neste modelo consistia de 53 empresas que havia entrado em falência e 58 que não havia entrado. Os autores buscaram dividir igualmente a amostra quanto à natureza da indústria (fabril e varejo) tanto para um grupo quanto

para o outro. Foram utilizadas inicialmente 27 variáveis, e os refinamentos estatísticos incorporados já puderam ser observados na própria definição das variáveis: algumas delas foram usadas na forma logarítmica a fim de evitar a influência de *outliers* e assim atender aos requisitos da técnica, como a normalidade da sua distribuição.

Por fim, o novo modelo definitivo foi reduzido a sete variáveis, quais sejam:

X₁, Indicador de Retorno sobre Ativo:

Medida do lucro antes dos juros e imposto de renda sobre o ativo total da empresa. Esta variável também havia se mostrado estatisticamente significativa no estudo para o *Z-Score* de 1968.

X₂, Indicador de estabilidade das receitas:

Medida aferida através do erro padrão da estimativa da variável X_1 por um período de dez anos. Essa variável capta a flutuação das receitas, sendo uma *proxy* para o risco do negócio.

X₃, Serviço da Dívida:

Medida pelo logaritmo do índice de cobertura de juros sobre o razão entre o capital de giro e o endividamento de longo prazo. Essa variável foi medida através do logaritmo na base 10 para melhorar a homocedasticidade da medida.

X₄, Lucratividade Acumulada:

Medida pelos lucros retidos sobre o ativo total. Essa variável, a mais significativa do modelo de 1968, embute aspectos como a idade e a política de dividendos da empresa.

X₅, Liquidez:

Medida através do índice de liquidez corrente.

X₆, Capitalização:

Medida pela razão entre o patrimônio líquido e o capital total. Os dois termos, tanto o numerador quanto o denominador, foram mensurados pelo seu valor de mercado e através da média dos últimos cinco anos. Os autores buscaram através dessa média incluir uma medida que acrescentasse um componente temporal além de suavizar possíveis flutuações severas.

X₇, Tamanho:

Medida pelo logaritmo do valor do ativo total. A transformação se deu para normalizar a variável.

Como esse modelo foi desenvolvido em parceria com uma empresa para fins comerciais, a equação final da análise discriminante não foi divulgada, e portanto estudos para testar o modelo não podem ser facilmente realizados quanto era com o modelo anterior. Comparando-se os dois modelos através do Quadro 1 abaixo, pode-se verificar que os resultados obtidos com o *Modelo ZETATM* são superiores aos resultados do *Modelo Z-Score*, principalmente nos anos 2 a 5 para empresas insolventes. O *Modelo Z-Score* apresenta resultados ligeiramente melhores nas previsões de empresas solventes.

ANOS ANTES DA INSOLVÊNCIA	ZETA™ MODEL		Z-SCORE	
	INSOLVENTE	SOLVENTE	INSOLVENTE	SOLVENTE
1	96,2	89,7	93,9	97,0
2	84,9	93,1	71,9	93,9
3	74,5	91,4	48,3	N.D.
4	68,1	89,5	28,6	N.D.
5	69,8	82,1	36,0	N.D.

Tabela 8: Quadro comparativo entre a precisão de previsão do Modelo ZETA™ e o Modelo Z-Score em percentagens. Fonte: adaptado de Altman (1983, p.137).

O teste realizado pelos autores para saber se as matrizes de variância-covariância são diferentes deu significativo, ou seja, apontou para a hipótese alternativa de que as matrizes são diferentes. Assim, uma estrutura quadrática deveria ser usada no modelo. No entanto um dilema foi criado quando eles perceberam que a precisão do modelo era maior usando-se uma equação linear do que uma equação quadrática. Isso fez com que eles optassem por usar a forma linear ao invés da quadrática, apesar da recomendação estatística afirmar o contrário.

2.4.2.3 Modelo de Altman, Baidya e Dias (1979)

Em 1979 foi publicado o primeiro estudo brasileiro a utilizar a metodologia apresentada por Altman para resolver a questão da previsão de problemas financeiros em empresas. A situação econômica brasileira da época era de “altas - e persistentes - taxas de inflação” (Altman, Baidya e Dias, 1979, p.18). As taxas de crescimento da economia brasileira à época era cerca de 5% ao ano; taxas consideradas boas mesmo ao nível internacional. Porém, as empresas passavam por uma situação de endividamentos altos, cuja consequência imediata eram as crescentes despesas

financeiras, implicando em saídas de caixa sujeitas ainda às correções monetárias. Os dados também mostravam que de 1972 a 1976 o número de concordatas e falências requeridas no Rio de Janeiro e em São Paulo havia caído, mas que a tendência se inverteu no período seguinte, crescendo 55% e 28% respectivamente entre 1976 e 1977¹².

Na realização da pesquisa, os autores identificaram 23 empresas com reconhecidos problemas financeiros e outras 35 empresas sem problemas financeiros aparentes. Foi tomado o cuidado de buscar, na amostra de controle, empresas do mesmo setor de atividade e também de porte semelhante. Os autores buscaram ainda, sempre que possível, escolher duas empresas do mesmo setor e do mesmo porte para cada empresa com problemas financeiros¹³.

A classificação das empresas nos dois grupos (empresas com problemas e empresas sem problemas) foi realizada adotando-se um critério objetivo e bem definido. As empresas fariam parte do grupo de empresas com problemas financeiros se pudessem ser enquadradas em pelo menos um dos casos a seguir:

- a empresa teve um pedido formal de falência;
- a empresa teve um pedido formal de concordata;
- a empresa teve soluções extrajudiciais para reorganizar a operação ou a estrutura financeira; ou
- a empresa teve suas atividades encerradas sem recorrer a meios legais.

Na escolha das variáveis do estudo, Altman, Baidya e Dias não adotaram um grande número de variáveis como Altman havia realizado em seus dois estudos

¹² Dados referentes aos primeiros seis meses de cada um dos anos mencionados.

¹³ Os autores buscaram ainda, sempre que possível, escolher empresas privadas nacionais, assumindo que o risco de insolvência é menor para empresas estatais e empresas subsidiárias de multinacionais.

anteriormente descritos. Para este estudo com dados brasileiros, foram adotadas as mesmas variáveis já selecionadas do estudo de Altman de 1968. Apenas algumas modificações tiveram que ser feitas em duas das cinco variáveis para que se adequassem à realidade brasileira.

Na variável “lucros retidos sobre ativo total” (variável X_2), os autores se depararam com uma questão conceitual: nos Estados Unidos, os lucros retidos registram os lucros acumulados pela firma menos os dividendos em dinheiro pagos aos acionistas. No Brasil, a incorporação dos lucros ao capital com distribuição de bonificação em ações e as reservas criadas para corrigir a desvalorização inflacionária do capital de giro que são deduzidas dos lucros inviabilizam o uso do valor lucros retidos. Assim, como forma de buscar uma conta que fosse compatível com o conceito da conta americana, calculou-se a variável X_2 como sendo:

$$X_2 = \frac{\text{não exigível} - \text{capital aportado pelos acionistas}}{\text{ativo total}}$$

A outra variável modificada pelo modelo de Altman, Baidya e Dias foi a variável “Valor de Mercado do Patrimônio Líquido sobre Valor Escritural do Passivo Total” (X_4). Como são poucas as empresas brasileiras com ações negociadas na bolsa de valores, o cálculo do valor de mercado do Patrimônio Líquido ficou inviabilizado. Como alternativa, foi usado o valor contábil do Patrimônio Líquido, que foi então dividido pelo Exigível Total. A variável X_4 ficou sendo:

$$X_4 = \frac{\text{não exigível}}{\text{exigível total}}$$

Assim, as cinco variáveis usadas no estudo são as seguintes:

- X_1 , Capital de Giro Líquido/Ativo Total;
- X_2 , Não exigível menos o Capital Aportado pelo acionista/Ativo Total;

- X_3 , Lucro Líquido Antes de Juros e Imposto de Renda/Ativo Total;
- X_4 , Não exigível/Exigível Total;
- X_5 , Vendas/Ativo Total;

Ao executar a solução do modelo discriminante, os autores obtiveram dois modelos, cada um com quatro variáveis. O primeiro modelo, após um critério de seleção de variáveis adotado, incluía as variáveis X_2 , X_3 , X_4 e X_5 . O segundo modelo não incluiu a variável X_2 . Assim, os dois modelos discriminantes encontrados no estudo são os seguintes:

$$\begin{cases} Z_1 = -1,44 + 4,03X_2 + 2,25X_3 + 0,14X_4 + 0,42X_5 \\ Z_2 = -1,84 - 0,51X_1 + 6,32X_3 + 0,71X_4 + 0,52X_5 \end{cases}$$

Em ambos os modelos, o ponto de corte discriminante é zero, ou seja, valores de Z (tanto Z_1 como Z_2) abaixo de zero indicam empresas como tendo problemas financeiros, e valores acima de zero indicam empresas como não tendo problemas financeiros. Dos dois modelos encontrados, os autores verificaram que o modelo 1 possuía desempenho melhor para dados referentes aos penúltimos e antepenúltimos anos antes do problema. Usando-se portanto o modelo 1, constatou-se que a precisão global do modelo era de 88%, ou seja, 7 das 58 empresas foram mal classificadas. O erro tipo I que é o de classificar uma empresa com problema financeiro no grupo das sem problemas financeiros foi de 13% (3 em 23), e o erro tipo II que é o de classificar uma empresa sem problema financeiro no grupo das com problemas financeiros foi de 11,4% (4 em 35). Os resumos dos dados apresentados estão mostrados nas duas tabelas a seguir.

		Preditas		N
		Insolventes	Solventes	
Insolventes (Real)		20	3	
	Número de Predições Corretas	Percentual de Predições Corretas	Percentual de Erros	
Tipo I	20	87	13	23

Tabela 9: Precisão dos testes de classificação para empresas insolventes.

		Preditas		N
		Insolventes	Solventes	
Solventes (Real)		4	31	
	Número de Predições Corretas	Percentual de Predições Corretas	Percentual de Erros	
Tipo I	31	88,6	11,4	35

Tabela 10: Precisão dos testes de classificação para empresas solventes.

O intervalo para valores de Z entre $-0,34$ e $+0,20$ foi considerado a zona de ignorância, pois foi nesse intervalo onde ocorreram os erros de classificação. Nenhuma empresa com escore Z menor que $-0,34$ ou maior que $+0,20$ foi classificada erroneamente.

Como validação, três procedimentos foram adotados. O primeiro deles foi baseado no teste de Lachenbruch (1967), que adota um procedimento que retira uma das empresas da amostra e processa novamente o modelo para $n - 1$ casos. Após a geração de uma nova função discriminante, o caso retirado da amostra é classificado. O resultado dessas n classificações é computado para se saber a precisão global do modelo. Este método é conhecido como *leaving-one-out method* (Lachenbruch, 1975). A precisão obtida nesta validação foi de 87%, mesmo valor da precisão obtida no modelo de classificação para empresas insolventes.

Em seguida, como segunda metodologia para validação do modelo, foram extraídas cinco sub-amostras, cada uma com doze empresas com problemas

financeiros e dezoito sem problemas. Para cada uma dessas sub-amostras foi gerado um modelo discriminante. Com os modelos gerados, as vinte e oito empresas restantes foram classificadas e obteve-se a precisão global média de 84,7% para ambos os tipos de erro.

Por fim, foram usados os dados das empresas para um e dois anos antes da data de registro da situação de problema financeiro. Com um ano antes, foi possível classificar corretamente 84,2% das empresas com problemas, e para dois anos antes foi possível classificar corretamente 77,8% das empresas com problemas.

O resultado geral desse estudo pode ser classificado como bom, apesar das eventuais críticas que possam recair sobre a metodologia empregada e sobre a falibilidade das demonstrações financeiras publicamente disponíveis. Este, assim como os modelos anteriores, mostrou boa capacidade discriminatória das empresas em dois grupos distintos, e foi capaz de prever com relativa precisão o problema a que se propunha.

Na seção seguinte, serão examinados os modelos que usam a Regressão Logística como forma de acessar o mesmo problema. Esta técnica surgiu após as críticas sofridas pelos modelos discriminantes e apresentaram também resultados bastante promissores.

2.4.3 Modelos Usando a Regressão Logística

2.4.3.1 Modelo de Ohlson (1980)

Em 1980, James Ohlson desenvolveu uma estimativa de falência probabilística através de um modelo de Regressão Logística (RL). O modelo logístico foi adotado para estudos de problemas desta natureza em virtude principalmente da Análise Discriminante Multivariada (ADL) impor às variáveis independentes condições do tipo: serem normalmente distribuídas e terem suas matrizes de variância-covariância iguais entre os dois grupos de empresas (com e sem problemas financeiros). Outro aspecto importante e que é também motivo de crítica no método da ADL é que o resultado da expressão discriminante fornece um escore, que possui pouca interpretação intuitiva. Como afirma Ohlson (1980, p.112), este escore é basicamente um dispositivo (discriminatório) de classificação ordinal, não tendo embutido nenhum aspecto probabilístico na medida. A outra fonte de crítica imposta por Ohlson foi com relação aos critérios de emparelhamento utilizados até então pelos pesquisadores do método de ADL: na visão de Ohlson, os critérios adotados eram arbitrários e não se sabe ao certo quais os ganhos e perdas associados com o uso de diferentes critérios de emparelhamento.

O uso da RL evita todos estes problemas mencionados acima. Além de não depender da exigência de normalidade das variáveis independentes e da igualdade das matrizes de covariância, a RL é semelhante a uma regressão múltipla, pois possui o poder de incorporar efeitos não lineares (Hair e outros autores, 1998). Ohlson resume o problema de estimação após o uso da RL na seguinte frase:

dado que uma firma pertence a alguma população pré-especificada, qual é a probabilidade de que a firma venha a fracassar dentro de um período de tempo pré-especificado? (Ohlson, 1980, p.112).

Ohlson critica a omissão das pesquisas anteriores por não mencionarem um problema importante na montagem do banco de dados das empresas a serem analisadas: a questão da data de disponibilização dos dados financeiros das empresas. Uma empresa nunca publica seus dados de balanço logo após o encerramento do ano fiscal, pois eles inclusive ainda deverão passar por um parecer de auditores. Sendo assim, é possível saber se uma empresa entrou em concordata ou falência antes ou depois da data de publicação de seus demonstrativos financeiros. O argumento de Ohlson é que a avaliação da capacidade de previsão de um modelo requer que as variáveis independentes estejam disponíveis para uso antes do evento que se quer estudar. Se forem utilizados dados publicados após a consumação do fato, haverá uma sobrevalorização da capacidade de previsão do modelo encontrado. Para resolver esse problema e não incorrer no mesmo erro, Ohlson recorreu a uma outra fonte de dados, diferente da usada por seus predecessores onde naquela havia a data em que o demonstrativo havia sido disponibilizado para o público.

Para a seleção das empresas para compor cada um dos grupos de estudo, Ohlson usou o conceito legalista de dificuldade financeira. Ele colheu dados de empresas que entraram com pedido de concordata no período de 1970 a 1976. Os pedidos de concordata considerados foram tanto o Capítulo 10 como o Capítulo 11 e ainda outras formas de notificação indicativa de procedimentos de concordata. As restrições à seleção das empresas, além da data de ocorrência do evento, exigiam que elas não pertencessem ao setor de serviços, de transporte ou financeiros, e que tivessem tido

ações negociadas nos três anos anteriores ao pedido de concordata. Além disso, o outro fator limitante foi tão somente a disponibilidade de dados financeiros.

Foram usadas neste estudo 105 firmas concordatárias e 2058 firmas não concordatárias. Segundo Ohlson, as variáveis adotadas para compor a estimação logística foram definidas com base na simplicidade. Os modelos gerados possuíam um intercepto e nove indicadores, que estão listados abaixo:

- **Tamanho (Porte) da Firma**
- **TLTA** = Valor Total de Passivos dividido por Valor Total de Ativos;
- **WCTA** = Capital de Giro dividido por Valor Total de Ativos;
- **CLCA** = Passivo Circulante dividido por Ativo Circulante;
- **OENEG**: Assume o valor 1 se o Valor Total de Passivos exceder o Valor Total de Ativos, e 0 se não exceder;
- **NITA** = Lucro Líquido dividido por Valor Total de Ativos;
- **FUTL** = Fundos provenientes das operações, dividido pelo Valor Total do Passivo;
- **INTWO** = Assume o valor 1 se o Lucro Líquido for negativo nos dois últimos anos; do contrário, 0;
- **CHIN** = $(NI_t - NI_{t-1}) / (|NI_t| + |NI_{t-1}|)$, onde NI_t é o Lucro Líquido do período mais recente. Esta variável busca medir mudança no Lucro Líquido.

Ohlson aplicou três modelos em seu experimento. O primeiro foi capaz de prever corretamente 96,12% dos casos de falência com um ano de antecedência. O segundo, para dois anos de antecedência (dado que a empresa não irá à falência no ano subsequente), foi capaz de prever corretamente 95,55% dos casos. E para o terceiro modelo, que previa falência em qualquer dos dois anos seguintes, o percentual de acerto foi de 92,84%.

O modelo logístico permite calcular ainda o grau de perfeição (*goodness-of-fit*) com que o modelo ajusta os dados reais. Essa medida é semelhante ao R^2 conhecido

da regressão. Os resumos dos percentuais de predição e dos coeficientes para os 3 modelos são mostrados nas tabelas a seguir.

<i>MODELOS</i>	<i>ÍNDICE (GOODNESS-OF-FIT)</i>	<i>PERCENTUAL PREDITO CORRETAMENTE</i>
MODELO 1	0,8388	96,12
MODELO 2	0,7970	95,55
MODELO 3	0,7190	92,84

Tabela 11: Resultados dos Experimentos de Ohlson, 1980.

	<i>Variáveis</i>									
	<i>Tamanho</i>	<i>TLTA</i>	<i>WCTA</i>	<i>CLCA</i>	<i>NITA</i>	<i>FUTL</i>	<i>INTWO</i>	<i>OENEG</i>	<i>CHIN</i>	<i>CONST</i>
<i>Modelo 1</i>										
Estimativas	-0,407	6,03	-1,43	0,0757	-2,37	-1,83	0,285	-1,72	-0,521	-1,32
Estatística t	-3,78	6,61	-1,89	0,761	-1,85	-2,36	0,812	-2,450	-2,21	-0,970
<i>Modelo 2</i>										
Estimativas	-0,519	4,76	-1,71	-0,297	-2,74	-2,18	-0,780	-1,98	0,4218	1,84
Estatística t	-5,34	5,46	-1,78	-0,733	-1,80	-2,73	-1,92	-2,42	2,10	1,38
<i>Modelo 3</i>										
Estimativas	-0,478	5,29	-0,990	0,062	-4,62	-2,25	-0,521	-1,91	0,212	1,13
Estatística t	-6,23	7,72	-1,74	0,738	-3,60	-3,42	-1,73	-3,11	1,30	1,15

Tabela 12: Resultados dos modelos de regressão logística de Ohlson, 1980.

O percentual de empresas preditas corretamente pelos três métodos indicados foi sempre acima de 90%. No entanto, não é possível simplesmente declarar que os resultados obtidos foram bons. Isto porque se classificássemos todas as empresas do estudo como pertencentes ao grupo daquelas com problemas financeiros, 91,15% das empresas seriam classificadas corretamente¹⁴. Para se saber se o resultado alcançado

¹⁴ $\frac{2058}{105 + 2058} = 0,9115$

foi de fato bom ou ruim é necessário conhecer a probabilidade *a priori* de uma empresa se tornar concordatária. Assim, o ponto de corte para este estudo não pode ser 0,50, mas sim um valor que minimize os erros de classificação das empresas (erros Tipo I e Tipo II). Ohlson testou diversos pontos de corte para os modelos 1 e 2 a fim de encontrar aquele que minimiza a soma dos erros do Tipo I e II. Para o modelo 1, o ponto de corte deveria ser 0,038. Para este valor, 17,4% das empresas saudáveis e 12,4% das empresas não saudáveis são classificadas incorretamente. Para todos os outros pontos de corte adotados, a soma desses dois valores encontrados seria sempre maior do que 29,8% (17,4% + 12,4%). Para o modelo 2, o ponto de corte deveria ser de 0,08. Para este valor de corte, o erro do Tipo I seria de 20,2% e o erro de Tipo II seria de 8,6%. A soma dos dois erros seria portanto de 28,8%, um pouco menor que o erro do modelo 1.

A comparação deste estudo de Ohlson com os estudos anteriores que usaram a técnica de ADL era impossível de ser feita. Primeiro porque havia o problema não atentado nos estudos anteriores quanto a não usar demonstrações financeiras que pudessem ter sido divulgadas após a decretação da situação de falência/concordata. Isso fez com que o tempo médio entre a data de publicação do balanço e a data do pedido de falência/concordata fosse maior para este estudo do que para os estudos anteriores. Em segundo lugar, havia poucos estudos realizados com dados deste período da década de 1970, e comparar épocas distintas não é adequado para estudos desta natureza.

2.4.4 *Uso de Redes Neurais Artificiais em Previsão de Falência*

Na definição de Klimasauskas (1996), Redes Neurais Artificiais (RNA) são uma tecnologia de processamento de informação inspirada por estudos do cérebro e do sistema nervoso humano. O uso das RNA no tratamento de dados para análise de falência ainda é recente. Todavia, já há alguns trabalhos publicados que servem de suporte a uma análise de sua utilidade. Segundo Almeida & Dumontier rede neural é

(...) um instrumento que permite às empresas explorarem seu ambiente de maneira privilegiada e assegurarem sua perenidade, transformando dados em informações úteis à tomada de decisão, aumentando o valor agregado de seus produtos e serviços (1996, p.53).

O estudo das RNA é originário do esforço em simular o funcionamento do cérebro humano (obviamente que em uma escala muito menos complexa). Elas são especialmente úteis para simular inteligência em atividades como reconhecimento de padrões, em atividades de associação e classificação. E é exatamente por esta sua capacidade de atuar como ferramenta de classificação que as RNA são úteis na solução de problemas como o que esta dissertação trata.

As RNA não devem ser confundidas com os sistemas especialistas (*expert systems*). Estes sistemas são criados a partir da experiência e com base no conhecimento de um especialista e na maneira como ele toma suas decisões. A grande questão com que se deparam os sistemas especialistas é como inserir em um conjunto de regras aspectos inerentes ao conhecimento, intuição e sensibilidade humana a fim de tomar a melhor decisão. Como afirma Almeida (1995), as RNA não precisam de um especialista para a criação de uma base de conhecimentos - ele é adquirido a partir de uma base de dados.

Segundo Trippi e Lee (1996), as RNA são formadas por diversos processadores, todos programados para executar uma mesma tarefa. Cada processador desses é chamado de neurônio ou nó, onde cada neurônio possui uma única saída e pode possuir uma ou mais entradas, como pode ser visto na figura abaixo.

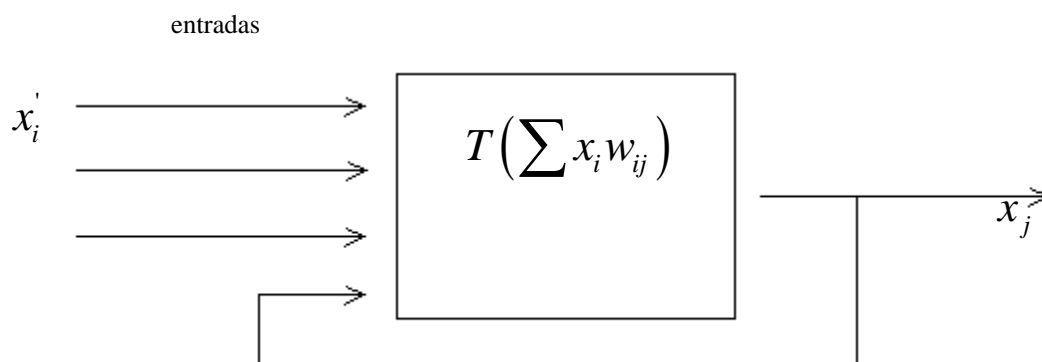


Figura 4: Processadores das Redes Neurais Artificiais.

As saídas dos neurônios podem ser as entradas de um outro neurônio, formando uma verdadeira rede, até que os últimos neurônios produzem a saída final da rede. O neurônio artificial (análogo ao neurônio biológico) recebe sinais de entrada (análogos aos impulsos eletroquímicos) de outros neurônios. A saída do neurônio artificial corresponde aos sinais enviados de um neurônio biológico através de seu axônio. Esses sinais artificiais podem mudar de forma similar às mudanças que ocorrem nas sinapses da rede neural humana (MEDSKER, TURBAN & TRIPPI, 1993, pp.5-6).

Em geral, os processamentos que ocorrem em cada neurônio são extremamente simples. Na maioria das vezes toma-se uma soma ponderada de entradas e calcula-se um valor de saída, resultado da função do tipo:

$$x_j = T\left(\sum_i w_{ij} x_i\right),$$

onde x_j é o elemento de saída, w_{ij} é o coeficiente de ponderação entre os elementos i e j e T é uma função de transferência, também chamada de função transformação. As funções de transferência mais comuns são as do tipo sigmoideal e suas variações:

$$T(y) = \left(1 + e^{-y}\right)$$

A vantagem de uma transformação assim é que os valores de saída ficam padronizados (por exemplo, entre zero e um), evitando que após inúmeros cálculos o número se torne muito grande.

A grande característica de distinção entre os diferentes tipos de RNA é a forma de interconexão entre os nós. As redes podem ainda ser combinadas em camadas que operam em sincronia umas com as outras, existindo uma camada de entrada, uma camada de saída e uma ou mais camadas intermediárias. As camadas intermediárias servem para expressar uma relação não-linear entre os nós de entrada e de saída.

As RNA são capazes de aprender através de exemplos não só com dados quantitativos como também com dados qualitativos, através de princípios associativos (ALMEIDA & DUMONTIER, 1996, p.53), uma vez que elas não impõem restrições para que as variáveis de previsão sejam separáveis linearmente nem que sejam independentes (COATS & FANT, 1996). Elas são capazes de produzir soluções a problemas em que os dados são incompletos e/ou distorcidos (ALMEIDA & DUMONTIER, 1996; HAWLEY, JOHNSON & RAINA, 1996).

A aprendizagem das redes neurais artificiais ocorre continuamente. O aprendizado se dá através de mudanças incrementais nos pesos dos coeficientes de interconexão. Essas mudanças ocorrem de acordo com uma das regras de aprendizado existentes.

Na construção da rede, os pesos são atribuídos às conexões entre os neurônios, e os valores de entrada são transformados ao passarem de neurônio a outro (ou de camada a outra). Esses pesos são geralmente arbitrados no início do processo e vão sendo corrigidos à medida que o aprendizado evolui com a utilização dos casos representativos do problema em estudo (ALMEIDA & DUMONTIER, 1996, p.54). Quando todos os casos forem aplicados, o processo inteiro é repetido até que o grau de redução do erro não seja mais significativo. Essas passagens são um ponto crítico no processo de aprendizagem de uma RNA. Podem ser necessárias milhares de passagens até que a rede fique estável, e o tempo necessário para o aprendizado é um aspecto importante a ser observado, uma vez que o processamento pode se tornar extremamente demorado. Depois de atingido o ponto de estabilidade da rede, ela pode ser usada para fazer a classificação de novos casos. Este novo processamento é então muito mais rápido do que o processamento na fase de aprendizagem, pois requer apenas uma passagem dos casos pela rede.

Os modelos de RNA pecam por não possuírem capacidade explanatória, ou seja, a saída final de um modelo baseado numa RNA raramente contribuirá para formar uma teoria (TRIPPI & LEE, 1996). O caráter de sucesso ou fracasso de uma RNA é medido estritamente com base em resultados alcançados, e por isto mesmo é sempre útil comparar os resultados obtidos com os resultados de um modelo estatístico como uma regressão, por exemplo.

Além da já mencionada capacidade de tratar também dados qualitativos, as RNA não precisam atender pressupostos quanto à distribuição do conjunto de dados amostrais e também são menos sensíveis ao efeito provocado pelo tamanho da

amostra. A multicolinearidade também é um problema de conseqüências menores, mas é sempre útil eliminar variáveis altamente correlacionadas. A seleção de variáveis por meio de uma regressão *stepwise* é uma maneira de tentar tornar o modelo mais eficiente do que usando todo o conjunto de variáveis (TRIPPI & LEE, 1996).

A aplicação de um modelo de RNA não pode garantir que a solução encontrada seja a melhor solução possível para determinado problema e conjunto de dados. Às vezes, até mesmo a repetição dos resultados para um mesmo conjunto de dados não pode ser alcançada. A grande crítica que se faz ao uso de redes neurais em Finanças como um modelo de previsão de falência é que não se sabe dizer como cada um dos fatores (indicadores contábeis) escolhidos para o modelo contribui para a resposta apresentada. Além disso, os pesos entre as conexões não podem ser interpretados de maneira tão óbvia quanto é possível em uma análise discriminante (MEDSKER, TURBAN & TRIPPI, 1993, pp.11-3).

Como afirmam Hawley, Johnson & Raina (1996), a estrutura interna das RNA tornam difícil seguir os passos usados para alcançar o resultado final, tornando o sistema uma espécie de caixa obscura.

O processo de aprendizagem das RNA se dá de forma que os neurônios sejam treinados a reconhecer os padrões de classificação através dos dados de entrada do problema. Segundo Medsker, Turban & Trippi (1996, pp.16), o procedimento consiste em apresentar para o neurônio a seqüência dos padrões de entrada de forma que os pesos das conexões sejam ajustados pelo computador após cada interação. Esta operação é repetida até que os pesos convirjam para um conjunto de valores que permita ao neurônio classificar corretamente cada um dos dados de entrada.

2.4.4.1 *Estudo de Odom e Sharda (1993)*

Um estudo comparativo entre os modelos discriminante e as redes neurais foi realizado por Odom & Sharda (1993, pp.177-85). Os resultados obtidos nessa pesquisa mostram evidências de que as RNA são modelos eficientes de predição de falência. Além disso, elas se mostraram mais robustas do que a Análise Discriminante para amostras de tamanho reduzido.

Neste estudo os autores usaram o estudo de Altman de 1968 como base de comparação para seus resultados. Por esta razão, eles usaram inclusive as mesmas variáveis usadas por Altman¹⁵. A RNA usada neste estudo para a etapa de aprendizagem foi uma rede *back-propagation* com uma camada de entrada, uma camada intermediária e uma camada de saída. As camadas de entrada e intermediária possuíam cada uma cinco neurônios. A camada de saída possuía apenas um neurônio, que fornecia como resultado de saída um valor entre 0 ou 1, representando uma empresa insolvente ou solvente respectivamente. Os dados disponíveis dos cinco indicadores do estudo foram apresentados à rede, que conseguiu na fase de aprendizagem classificar corretamente 100% das empresas. A convergência do processo demorou aproximadamente 24 horas, após o total de 191.400 iterações.

Os autores neste estudo fizeram uma avaliação da robustez das técnicas ao ajustar as amostras de treinamento às proporções em que empresas solventes e insolventes são geralmente encontradas no mundo real. Eles então formaram 3

¹⁵ Ver seção 2.4.2.1 para rever o estudo de Altman (1968).

grupos, um com proporção 50/50, outro com 80/20 e outro com 90/10. Os resultados dos testes são apresentados na tabela abaixo.

Proporções Adotadas		50/50		80/20		90/10	
Modelo	Real	Insolvente	Solvente	Insolvente	Solvente	Insolvente	Solvente
	Predito	(27)	(28)	(27)	(28)	(27)	(28)
Redes Neurais Artificiais	Insolvente	22 (81,5)	5 (18,5)	21 (77,8)	6 (22,2)	21 (77,8)	6 (22,2)
	Solvente	5 (17,9)	23 (82,1)	6 (21,4)	22 (78,6)	4 (14,3)	24 (85,7)
Análise Discrimina nte	Insolvente	16 (59,3)	11 (40,7)	19 (70,4)	8 (29,6)	16 (59,3)	11 (40,7)
	Solvente	3 (10,7)	25 (89,3)	4 (14,3)	24 (85,7)	6 (21,4)	22 (78,6)

Tabela 13: Comparação dos resultados obtidos com a Análise Discriminante e Redes Neurais.

Fonte: Odom & Sharda (1993, pp.182).

A análise da tabela acima mostra que as RNA apresentam resultados mais robustos do que a ADL, qualquer que seja a proporção empregada. Uma análise ainda mais aprofundada dos resultados de cada caso mostrou que as empresas classificadas incorretamente pela RNA também foram um problema para a ADL.

2.4.4.2 Estudo de Altman, Marco & Varetto (1994)

Usando uma extensa base de dados da *Centrale dei Bilance*, uma organização formada pelo Banco da Itália, a Associação Bancária Italiana e outros 40 bancos e instituições de crédito, Altman, Marco & Varetto (1994) apresentam os resultados alcançados em um estudo onde se comparou a eficiência de dois modelos de previsão de dificuldades financeiras: o primeiro usando ADL e o segundo usando RNA.

Este estudo talvez não tivesse tanta importância se não tivesse sido escrito por um dos pioneiros de estudos de previsão de dificuldades financeiras : Edward Altman. Seus primeiros estudos (Altman, 1968, 1977, 1979) são referências obrigatórias quando se estuda este assunto, e vê-lo utilizar técnicas mais modernas acentuam o caráter evolutivo da ciência.

Usando mais de 1000 empresas italianas de médio porte e pequeno porte¹⁶ divididas em 2 grupos de acordo com sua situação financeira, os autores usaram um modelo de RNA com arquitetura de retropropagação. Na primeira fase, usando os dados de três anos antes da identificação de situação de dificuldades financeiras, buscou-se avaliar a eficácia dos valores numéricos da análise discriminante. Se a RNA pudesse reproduzir os resultados alcançados com um conjunto menor de indicadores, isto seria uma clara capacidade de adaptação e simplificação por parte da rede sobre a ADL. Os melhores resultados foram obtidos usando-se uma rede de duas camadas intermediárias (a primeira com 10 neurônios e a segunda com 4 neurônios) e uma camada de saída com apenas um neurônio.

Como entrada de dados os autores usaram 10 indicadores financeiros: 4 índices relativos à estrutura de capital e endividamento, 2 relativas a liquidez e 4 representando a lucratividade das empresas. Todos os neurônios estavam interconectados, ou seja, conectavam-se com todos os outros neurônios do nível seguinte, incluindo os sinais de entrada, que estavam conectados a todos os neurônios da primeira camada. Foram usados 1000 ciclos na etapa de aprendizagem da rede, e os autores acreditaram que se esta fase tivesse sido um pouco mais longa o erro

¹⁶ Foram selecionadas apenas empresas cujas vendas anuais são ultrapassassem US\$ 60 milhões de dólares.

poderia ter sido reduzido ainda mais, o que provou a tese investigada. Uma série de outros testes e avaliações de diferentes estruturas foram realizadas, e as principais conclusões do trabalho realçadas pelos autores foram as seguintes:

- As RNA foram capazes de superar o desempenho da ADL até mesmo quando foram utilizados um conjunto diferente e menor de indicadores;
- O uso de RNA está limitado ao longo tempo requerido no processamento da fase de aprendizagem e na necessidade de se executar diferentes testes até identificar a melhor estrutura da rede, além do cuidado a ser tomado com a armadilha do superajustamento;
- Os pesos resultantes inerentes ao sistema não são transparentes ao usuário da informação, e a possibilidade de derivar uma rede de comportamento ilógico não pode ser descartada;

Por estes motivos, os autores não destacam as RNA como uma técnica dominante sobre as técnicas estatísticas tradicionais (como a ADL). Para eles, os dois métodos devem ser usados em conjunto, visando extrair de cada um deles aquilo que têm de melhor, servindo um como técnica de validação do outro.

2.4.4.3 Estudo de Poddig (1995)

Realizado em cooperação com uma instituição financeira alemã e com o intuito de comparar a ADL com a RNA, o estudo de Poddig (1995) utilizou 150 empresas francesas que haviam ido à falência. Os dados eram de um período que ia de 1 a 3 anos antes da data de falência. A amostra de firmas saudáveis também possuía 150 empresas, escolhidas aleatoriamente durante o mesmo período.

Poddig usou 45 indicadores financeiros extraídos de demonstrações contábeis das empresas. Estes índices, selecionados pelo banco interessado no estudo, foram

escolhidos com base na utilização prévia dos mesmos em análises feitas pela instituição.

A amostra total de 150 firmas para cada grupo foi dividida em duas sub-amostras. A primeira sub-amostra, que foi chamada de amostra de treinamento continha 100 empresas, enquanto que as 50 empresas restantes formaram a amostra de classificação.

Usando dados dos 3 anos anteriores ao pedido de falência, foram construídos dois modelos discriminantes usando ADL¹⁷. O método *stepwise* empregado na seleção de variáveis gerou modelos com apenas 3 variáveis cada um. O percentual de previsões corretas de ambos os modelos estão mostrados na tabela abaixo.

MODELO 1			MODELO 2		
<i>Anos antes da falência</i>	<i>Amostra de treinamento</i>	<i>Amostra de classificação</i>	<i>Anos antes da falência</i>	<i>Amostra de treinamento</i>	<i>Amostra de classificação</i>
3	70,5%	78,0%	3	68,5%	80,0%
2	78,0%	81,0%	2	81,0%	88,0%
1	83,0%	86,0%	1	84,0%	90,0%
Média	77,2%	81,7%	Média	77,8%	86,0%

Tabela 14: Resultados da ADL para o estudo de Poddig (1995).

Por conveniência do autor e também por dificuldade de calcular separadamente o custo de incorrer no erro do Tipo I e no erro do Tipo II, o autor tratou os dois erros como iguais. Infelizmente ele não publicou separadamente os dois erros, mas declarou que foram bastante equilibrados entre si. Outro inconveniente no teste foi que o autor

¹⁷ A razão de terem sido construídos dois modelos foi por que os técnicos do banco interessado montaram um modelo, enquanto que o autor montou um outro modelo. Maiores detalhes quanto às técnicas empregadas não foram fornecido ao longo do texto original.

não declarou explicitamente quais foram as diferenças entre os dois modelos discriminantes nem quais as variáveis finais de cada um dos modelos obtidos.

Neste estudo foram usadas duas arquiteturas de RNA. A primeira delas foi do tipo retropropagação, que é uma das mais largamente utilizadas em aplicações na área financeira. A segunda delas foi uma versão modificada do *Learning Vector Quantiser* (LVQ).

Para a primeira arquitetura do estudo, a princípio dois métodos de ajustamento de pesos de interconexão seriam empregados: o primeiro foi com base no custo dos pesos (*weight cost*) e o segundo com base no decaimento dos pesos (*weight decay*). Na medida em que realizava seu estudo, Poddig ficou sabendo dos resultados bem sucedidos de uma nova técnica de cálculo dos valores dos pesos. Ele então implementou este novo método e refez parte de seu estudo, obtendo de fato resultados melhores.

A aplicação desta nova técnica de ajustamento de pesos exigia uma validação permanente da amostra de treinamento a fim de obter o ponto ótimo de parada do ajustamento. Poddig teve que dividir a amostra de treinamento, obtendo uma sub-amostra de validação. O processo de ajustamento não pararia até que a rede apresentasse o maior número de classificações corretas nesta sub-amostra de validação.

Outro aspecto importante a ser ressaltado no estudo foi que para o treinamento da rede as variáveis passaram por um procedimento visando padronizar as medidas. Duas formas foram empregadas: a primeira delas utilizava não somente um procedimento de padronização, mas também uma transformação não-linear, enquanto

que a segunda forma fazia um re-escalamento de cada variável dentro de um intervalo único padronizado entre o mínimo e o máximo de cada medida.

Os resultados encontrados quando se usou o método de ajustamento pelo custo do peso e a padronização em conjunto com a transformação não-linear estão mostrados na tabela abaixo.

MODELO RNA 1		
<i>Anos antes da falência</i>	<i>Amostra de treinamento</i>	<i>Amostra de classificação</i>
3	83,5%	86,0%
2	85,5%	83,0%
1	91,5%	92,0%
Média	86,8%	87,0%

Tabela 15: Resultado do modelo de RNA de Poddig (1995) usando ajustamento de pesos pelo custo e padronização em conjunto com transformação não-linear das variáveis.

Para o procedimento de ajuste estatístico dos pesos, os melhores resultados foram alcançados quando se usou a otimização manual da rede. Os resultados para este modelo estão mostrados na tabela abaixo.

MODELO RNA 2		
<i>Anos antes da falência</i>	<i>Amostra de treinamento</i>	<i>Amostra de classificação</i>
3	81,0%	84,0%
2	85,5%	89,0%
1	91,0%	93,0%
Média	85,8%	88,7%

Tabela 16: Resultado do modelo de RNA de Poddig (1995) usando ajuste estatístico dos pesos e otimização manual da rede.

Foram usadas entre 10 e 13 variáveis nos modelo de RNA. No entanto, o autor não menciona quais as variáveis empregadas nem como se deu o processo de seleção.

Os resultados pelo método de RNA foram superiores aos da ADL. Quanto ao uso da segunda arquitetura de RNA empregada, os resultados obtidos foram inferiores aos da ADL, e portanto não serão discutidos nem apresentados aqui.

2.4.4.4 Estudo de Almeida e Dumontier (1996)

Diferenciando-se da maioria dos estudos que fazem uso de RNA em problemas na área de Finanças, Almeida & Dumontier (1996) propõem uma estruturação dos métodos de aplicação, fugindo da característica comum que é a de basear-se na sensibilidade e experiência do pesquisador no processo de construção da rede. Utilizando um modelo de retropropagação e a Regra Delta Generalizada como regra de aprendizagem, os autores utilizaram 2.414 empresas do setor de transporte de carga rodoviário francês como amostra.

A abordagem estruturada proposta pelos autores envolve a concepção e execução de um plano experimental. Este plano visa explorar como cada parâmetro a ser escolhido influencia o desempenho da rede, buscando uma arquitetura mais simples e ao mesmo tempo mais eficaz. Após uma análise gráfica para avaliação do plano, as redes de melhor desempenho são introduzidas em um portfólio de redes. Segundo os autores, “a utilização de um portfólio de redes permite atenuar o erro produzido pelo uso de apenas uma rede” (p.55).

Após gerar 560 redes no total, com 56 estruturas diferentes, os autores chegaram a algumas conclusões importantes:

- As redes que usaram os valores dos indicadores por si mesmos tiveram desempenho inferior ao das redes que usaram a diferença entre os valores destes indicadores;
- Redes com 5 ou 10 neurônios tiveram desempenho inferior às outras redes utilizadas, com 40 e 80 neurônios;
- Não foi possível chegar a uma conclusão definitiva quanto ao número de camadas: em certos momentos uma rede com uma camada apresentava desempenho melhor, em outros os melhores resultados eram obtidos com duas camadas;
- As redes não apresentaram variações consideráveis de desempenho quando o conjunto de indicadores era modificado;
- Ao comparar o desempenho das RNA com um modelo de regressão logística, os autores constataram que as previsões feitas com as redes neurais são mais confiáveis do que as feitas com o modelo logístico.

3. Metodologia

A presente dissertação tem por objetivo desenvolver três modelos diferentes para acessar o problema da previsão de empresas em situação de dificuldades financeiras. Ela se propõe também a analisar qual destes métodos leva a resultados mais confiáveis nas previsões e que variáveis são estatisticamente contributivas para cada um dos modelos.

Foi portanto realizada uma pesquisa descritiva e quantitativa, utilizando-se técnicas que permitam inferir para a população objeto deste estudo os resultados encontrados a partir dos dados da amostra. Os sujeitos são empresas brasileiras não financeiras de capital aberto com dados de balanço publicados entre os anos de 1996 e 2002 e disponíveis no banco de dados da Comissão de Valores Mobiliários - CVM e Bolsa de Valores do Estado de São Paulo - BOVESPA. Os dados pesquisados foram obtidos a partir da divulgação realizada por estas duas entidades em seus respectivos sítios eletrônicos. A opção por empresas não financeiras foi feita por que as empresas financeiras possuem características muito particulares em seus balanços, fazendo com que difiram consideravelmente das empresas de outros setores.

As empresas foram divididas em dois grupos, distintos entre si por uma característica objeto da análise: estarem em situação de dificuldades financeiras. Como definição para este estudo, foi considerada como em dificuldade financeira a empresa que estiver em processo de liquidação, em situação de concordata ou falência dentro do período compreendido pelo estudo. Para compor o grupo de empresas solventes foram selecionadas aquelas que estão listadas pela Bolsa de Valores de São

Paulo - BOVESPA dentro de quaisquer dos dois níveis de governança existentes. A escolha por este critério se deu por acreditar que empresas que tem bons níveis de governança são teoricamente empresas sólidas e que podem ser consideradas o oposto de empresas em situação de dificuldades financeiras.

3.1. Definição Operacional das Variáveis

Devido à falta de uma teoria consolidada que forneça o completo embasamento do problema e a conseqüente escolha das melhores variáveis para o estudo de previsão, foram usadas inicialmente como variáveis independentes aquelas mais comuns em estudos que precederam este e que são largamente discutidas e utilizadas na literatura relativa ao tema, seja por sua comprovada importância ou por sua simplicidade conceitual. Silva (2001) e Matarazzo (1998) trazem uma compilação destas variáveis, que foram adotadas nesta pesquisa. Algumas dessas variáveis foram pouco úteis quanto à sua capacidade de contribuir para o poder de previsão do modelo. Neste caso, essas variáveis foram retiradas do modelo final e sua exclusão será devidamente justificada oportunamente.

As variáveis selecionadas inicialmente para compor o modelo estão listadas a seguir:

3.1.1 *Variáveis de Estrutura de Capital e Endividamento*

As variáveis que compõem o grupo de indicadores de estrutura de capital e endividamento são todas do tipo quanto menor, melhor, e portanto espera-se que seus coeficientes, caso sejam incluídas no modelo final, tenham sinais negativos.

- X_{01} : Imobilização do Patrimônio Líquido: é a razão entre o ativo permanente e o patrimônio líquido, indicando quanto do patrimônio líquido da empresa está aplicado no ativo permanente. Se a empresa aplica uma grande parte de seu patrimônio líquido no ativo permanente, sobrarão menos recursos para serem aplicados no ativo circulante e a empresa conseqüentemente será mais dependente de recursos de terceiros, o que aumenta seu risco empresarial.

- X_{02} : Participação de Capitais de Terceiros: é a razão entre capitais de terceiros (passivo circulante + exigível de longo prazo) e o patrimônio líquido da empresa, representando a dependência da empresa com relação a recursos externos.

- X_{03} : Composição do Endividamento: é a razão entre passivo circulante e capitais de terceiros, indicando quanto da dívida total deverá ser quitada no curto prazo. É a relação entre as dívidas de curto prazo e as obrigações totais da empresa.

3.1.2 *Variáveis de Liquidez*

As variáveis que compõem o grupo de indicadores de liquidez são todas do tipo quanto maior, melhor, e portanto espera-se que seus coeficientes, caso sejam incluídas no modelo final, tenham sinais positivos.

- X_{04} : Liquidez Geral: é a razão entre ativo circulante mais realizável de longo prazo e passivo circulante mais exigível de longo prazo. Representa quanto a empresa possui de recursos para quitar o total de suas dívidas.

- X_{05} : Liquidez Corrente: é a razão entre o ativo circulante e o passivo circulante. Representa o quanto a empresa tem em recursos correntes para quitar suas dívidas de curto prazo.

- X_{06} : Liquidez Seca: é a razão entre a soma das disponibilidades com títulos a receber e outros recursos de rápida conversibilidade em dinheiro (excluídos estoques) e o passivo circulante. Representa o quanto a empresa tem em recursos de alta conversibilidade em dinheiro para quitar suas dívidas de curto prazo.

3.1.3 *Variáveis de Rentabilidade*

As variáveis que compõem o grupo de indicadores de rentabilidade são todas do tipo quanto maior, melhor, e portanto espera-se que seus coeficientes, caso sejam incluídas no modelo final, tenham sinais positivos.

- X_{07} : Giro do Ativo: é a relação entre as vendas líquidas da empresa e o total dos ativos, que representam o total de investimentos efetuados na empresa, indicando o nível de eficiência com que os recursos da empresa são investidos.

- X_{08} : Retorno sobre Vendas: é a relação entre o lucro líquido da empresa e suas vendas líquidas no período, ou seja, o percentual de lucro da empresa em relação ao seu faturamento líquido. Este indicador também é conhecido como Margem Líquida.

- X_{09} : Retorno sobre o Ativo: é a relação entre o lucro líquido da empresa e seu ativo total, indicando a lucratividade da empresa em relação aos seus investimentos totais. Este indicador também é conhecido como Rentabilidade do Ativo.

- X_{10} : Retorno sobre o Patrimônio Líquido: é a relação entre o lucro líquido da empresa e o patrimônio líquido, indicando quanto a empresa obtém de lucro com relação ao capital próprio investido pelo acionista.

- X_{11} : Capital de Giro Líquido sobre os Ativos: é a razão entre o capital de giro líquido e os ativos totais da empresa, indicando quanto ela possui de capital de giro para cada real do ativo total.

- X_{12} : Lucros Retidos sobre os Ativos: é a razão entre os lucros retidos e os ativos totais da empresa, indicando quanto ela reteve em lucros para cada real do ativo total.

3.1.4 *Variáveis de Atividade e Rotação*

- X_{13} : rotação de Estoques: é definida como sendo a razão entre os estoques e o custo de mercadorias vendidas. A questão dos estoques em uma empresa deve ser abordada sob os seguintes aspectos: Primeiramente, quando as empresas vão aos bancos buscar financiamentos, são os estoques em geral que são oferecidos como garantia da dívida pretendida, pois representam o potencial de geração de recursos em dinheiro. Em segundo lugar, os estoques são aplicações de recursos em ativos circulantes, e quanto maior forem os estoques, maior será o comprometimento dos recursos da empresa. Sendo assim, esta variável será considerada como do tipo quanto

menor, melhor, e portanto espera-se que seu coeficiente, caso seja incluída no modelo final, tenha sinal negativo.

- X_{14} : Recebimento de Vendas: representa a razão entre as duplicatas a receber e as vendas líquidas. Esta variável é do tipo quanto menor, melhor, e portanto espera-se que seu coeficiente, caso seja incluída no modelo final, tenha sinal negativo.

- X_{15} : Pagamento de Contas: representa a razão entre as contas fornecedores e compras. Esta variável é do tipo quanto maior, melhor, e portanto espera-se que seu coeficiente, caso seja incluída no modelo final, tenha sinal positivo.

- X_{16} : Estoque sobre Venda: é a razão entre o valor em estoques e as vendas líquidas da empresa. Como se espera que os estoques não sejam altos, e que o volume de vendas seja tanto maior quanto possível, este é uma variável do tipo quanto menor, melhor, e portanto espera-se que seu coeficiente, caso seja incluída no modelo final, tenha sinal negativo.

3.2. Amostragem

O corte no tempo realizado neste estudo, empregando dados de empresas que entraram em dificuldades financeiras nos últimos cinco anos já traz embutido em si um critério de seleção amostral (Barros, 2001, p.72). Os anos selecionados para o estudo visam escolher dados de empresas dentro de um período de estabilização da moeda e não sofrer influências dos períodos de alta inflação que caracterizou a economia brasileira em anos anteriores.

A grande dificuldade em estudos desta natureza tem sido a de obter de uma fonte confiável uma grande quantidade de dados necessários ao desenvolvimento de modelos de predição. Portanto, uma vez que a coleta dos dados não proporcionou um grande número de empresas não financeiras com dados disponíveis e que atendam a definição de dificuldades financeiras adotada neste estudo, a amostra de 40 empresas que foi aqui adotada não pode ser caracterizada como aleatória, mas de conveniência. No entanto, com intuito de atribuir algum caráter aleatório na seleção das empresas que constituem as amostras de treinamento e de validação dos modelos, foi adotado um critério de seleção baseando-se em números pseudo-aleatórios com o intuito de simular um sorteio das empresas que compõem estes dois sub-grupos.

3.3. Técnicas de Análise de Dados Empregadas

As técnicas empregadas neste estudo são de largo uso dentro das mais diversas áreas de pesquisas aplicadas em ciências sociais ou em demais áreas técnicas. Foram usadas as técnicas estatísticas de Análise Discriminante Linear, Regressão Logística e as Redes Neurais Artificiais.

Não são raras as vezes em que os pressupostos inerentes a cada uma dessas técnicas, em especial a ADL, sejam violados quando de suas aplicações em problemas como este aqui proposto. Convém, portanto que seja feito um esclarecimento a respeito de cada uma das técnicas, assim como dos testes estatísticos subjacentes à comparação entre os modelos gerados.

3.3.1 A Análise Discriminante Linear

Segundo Hair *et alii* (1998, p.244) a ADL é uma técnica estatística apropriada para o caso em que se dispõe de uma variável dependente nominal ou não métrica e um conjunto de variáveis independentes do tipo métrica.

A ADL é capaz de lidar com dois ou mais grupos. Esta técnica envolve a derivação de uma variável que é a combinação linear de duas ou mais variáveis independentes que melhor discriminam os grupos definidos a priori. Essa discriminação é atingida com a atribuição de pesos para as variáveis de forma a maximizar a variância entre os grupos em relação à variância dentro dos grupos. O resultado dessa combinação linear produz a chamada função discriminante, que é uma função que segue a seguinte forma:

$$Z_{jk} = a + W_1 X_{1k} + W_2 X_{2k} + \dots + W_n X_{nk}, \text{ onde:}$$

Z_{jk} = escore Z discriminante da função discriminante j para o objeto k

a = intercepto

W_i = peso discriminante para a variável independente i

X_{ik} = variável independente i para o objeto k

Hair *et alii* (1998, p.245) afirmam que a Análise Discriminante é uma técnica estatística apropriada para testar a hipótese nula de que as médias para um conjunto de variáveis independentes são iguais. Para isso, a análise discriminante multiplica cada variável independente por seu respectivo peso e soma essas parcelas, resultando no escore Z de discriminação para cada indivíduo da análise. A média dos indivíduos de um grupo fornece a média do grupo, também chamada de centróide. Para cada grupo envolvido na análise discriminante será obtido um centróide. Esses centróides

indicam a localização mais provável de um indivíduo de um grupo particular, e a comparação entre os centróides dos grupos permite dizer o quão afastado estão estes grupos um do outro ao longo da dimensão sob teste.

O teste para significância estatística da função discriminante é uma medida generalizada da distância entre os centróides dos grupos. Esse cômputo é feito através da comparação entre a distribuição do escores discriminantes para cada um dos grupos. Se a sobreposição das distribuições for pequena, a função discriminante separa os grupos adequadamente. Se a sobreposição for grande, a função discriminante não é uma boa opção para a separação dos grupos. A figura a seguir ilustra o que foi dito com relação à adequação da função discriminante: a figura à direita ilustra uma discriminação adequada, enquanto que a figura à esquerda apresenta uma discriminação inadequada, uma vez que a sobreposição das funções é bem grande.

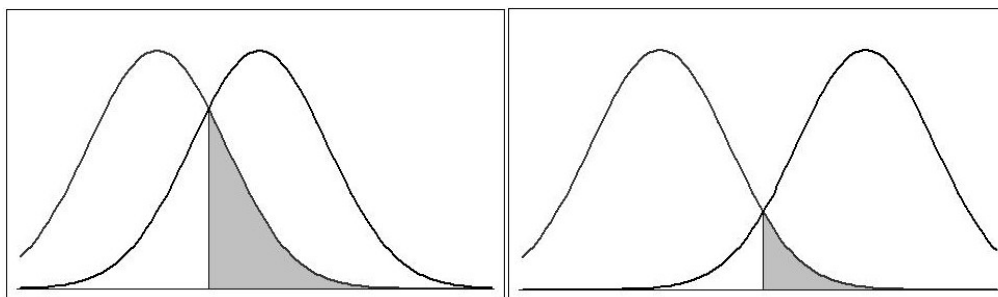


Figura 5: Representação univariada dos escores discriminantes.

Segundo Hair *et alii* (1998, p.255), a aplicação da análise discriminante pode ser vista a partir de uma perspectiva de construção de modelo em seis etapas, descritas nas subseções a seguir.

3.3.1.1 Problema de Pesquisa

Como é comum a todas as aplicações multivariadas, o problema de pesquisa tem destaque e é o primeiro estágio desse processo. Nesta etapa, os objetivos da pesquisa são ressaltados, a fim de esclarecer a natureza da aplicação do método. Neste estudo, a ADL tem por objetivos determinar quais variáveis independentes são responsáveis pela maior parte das diferenças entre os dois grupos de empresas (empresas solventes e insolventes) e desenvolver um modelo de classificação destas empresas em um dos dois grupos.

3.3.1.2 Modelo de Pesquisa

Em primeiro lugar, deve-se selecionar adequadamente as variáveis dependentes e independentes a serem aplicadas no modelo. Neste estudo, as empresas são classificadas em um de dois grupos mutuamente exclusivos e exaustivos: empresas solventes ou insolventes. Em outras palavras, cada empresa só pode ser classificada em um único grupo.

Em seguida, as variáveis independentes são definidas (ver seção 3.1 neste capítulo para a definição operacional das variáveis). Estas variáveis foram pré-selecionadas com base em pesquisas anteriores que abordaram este assunto e fazem parte do arcabouço teórico desenvolvido sobre o tema.

Hair et alli (1998, p.259) ressalta ainda a importância quanto ao tamanho da amostra: o tamanho do menor grupo deve exceder pelo menos o número de variáveis independentes selecionadas para o estudo, e cada grupo deve possuir pelo menos 20

observações. Os autores ressaltam também que uma amostra de validação seja adotada a fim de aferir se resultados confiáveis podem ser obtidos a partir do modelo desenvolvido. Uma vez que a quantidade de casos que esta pesquisa dispõe não é de todo suficiente para se fazer uma divisão entre amostra de análise e amostra de validação, a solução encontrada é a abordagem de *cross-validation* do tipo *leave-one-out*, onde a função é desenvolvida para $n - 1$ casos, e o caso n é utilizado para validação do modelo obtido. Este procedimento é repetido para todos os n casos da amostra. Esta abordagem de *cross-validation* produz um viés para cima nos resultados preditos pelo modelo, mas é sem dúvida uma melhor solução do que não validar de forma alguma os resultados alcançados.

3.3.1.3 *Pressupostos da ADL*

A fim de obter condições adequadas para o uso da técnica de ADL, uma série de pressupostos deve ser atendida pelo conjunto de empresas que compõem os casos em estudo. Em primeiro lugar, as variáveis independentes devem possuir uma distribuição aproximadamente normal. A fim de testar a hipótese nula de que as variáveis independentes pré-selecionadas para este estudo seguem distribuição aproximadamente normal os testes de Shapiro-Wilk, Anderson-Darling e de Kolmogorov-Smirnov com a correção de Lilliefors foram utilizados. Nos casos em que foram detectadas variáveis que não seguem uma distribuição aproximadamente normal, foi realizada uma das transformações de Box-Cox a fim de normalizar a distribuição da variável.

Em seguida, a igualdade das matrizes de variância-covariâncias dos dois grupos foi testada para a igualdade. Para se testar a hipótese nula desta igualdade, foi utilizado o teste M de Box. Matrizes de variância-covariância desiguais podem afetar negativamente o processo de classificação dos casos, pois os casos tendem a ser classificados dentro do grupo que apresenta a matriz de maior covariância. No caso da hipótese nula de igualdade das matrizes ser negada, Hair *et alii* (1998, p.259) sugere que sejam usadas para fins de classificação as matrizes específicas de cada um dos grupos.

Por fim, multicolinearidade das variáveis independentes pode afetar os resultados obtidos pelo modelo de classificação. Multicolinearidade quer dizer que uma variável é altamente correlacionada com uma ou mais variáveis, de forma que a primeira pode ser explicada pela combinação destas últimas.

A fim de identificar variáveis altamente correlacionadas e que pudessem trazer problemas para o modelo desenvolvido, foram conduzidos dois diagnósticos de multicolinearidade. O primeiro, feito através dos Fatores de Inflação de Variância (VIF - *Variance Inflation Factors*), mede o efeito de outras variáveis de previsão selecionadas pelo modelo sobre os coeficientes da regressão para uma determinada variável. A regra prática sugere que se adote como indicativo de existência de multicolinearidade os casos em que o VIF foi maior que 10 (Neter *et alii*, 1996; Hair *et alii*, 1998; Gujarati, 2000).

Uma outra análise a ser feita será através do Índice de Condição (CI - *Condition Index*) e da matriz de decomposição de variância dos coeficientes da regressão. O CI representa a colinearidade de combinações de variáveis. Como regra prática é adotado

que se o CI se situar entre 10 e 30, há colinearidade de moderada a forte¹⁸; se exceder 30, há multicolinearidade grave (Gujarati, 2000). Para todos os CI que ultrapassem o limite estabelecido, deve-se atentar para a matriz de decomposição de variância dos coeficientes de regressão. Proporções de variância acima de 90% indicam que há problemas de colinearidade, e ações para corrigir estes problemas devem ser adotadas.

Entre as medidas corretivas que podem ser adotadas em caso de detecção de multicolinearidade está a eliminação de uma (ou mais variáveis) colineares. Esta solução pode ser vista como uma das mais simples, mas corre-se o risco de estar cometendo um viés de especificação. Uma alternativa à eliminação de variáveis é a utilização de uma transformação nas variáveis afetadas, buscando eliminar ou reduzir a correlação existente (Gujarati, 2000).

Em seguida, um outro aspecto importante a ser detectado é a presença de casos de influência dentre aqueles em estudo. Hair *et alii* (1998, pp.221-37) descrevem detalhadamente procedimentos úteis ao tratamento de casos influentes. A princípio, o interesse deste estudo não é pela eliminação de eventuais casos influentes, mas apenas sua identificação e análise das classificações que sofreram. Algumas das empresas podem apresentar indicadores que a tornem diferentes das demais empresas da amostra, mas isto não quer necessariamente dizer que estes casos devam ser eliminados, pois empresas com tais diferenças existem no mundo real e não devem ser simplesmente eliminadas do modelo por simples comodidade, mas mantidas sob rigorosa observação.

¹⁸ Hair *et alii* (1998) adotam como limite mínimo para se detectar multicolinearidade o intervalo entre 15 e 30, ressaltando que 30 é o valor usualmente preferido.

3.3.1.4 Estimação da função Discriminante

O processo de estimação da função discriminante começa com a seleção das variáveis que compõem o modelo final. Existem algumas possibilidades que podem ser consideradas quanto ao critério de seleção destes indicadores, e entre elas se destacam principalmente o método *stepwise* e o procedimento de todas as regressões possíveis (*all-possible-regressions procedure*), que considera todos os subconjuntos de variáveis explanatórias existentes com o intuito de identificar um ou mais subconjuntos que possam ser considerados como bons modelos. Neter *et alii* (1996, pp.336-47) descrevem detalhadamente os procedimentos existentes a fim de se chegar ao(s) modelo(s) final(is). No caso de identificação de mais de um subconjunto de variáveis explanatórias, optou-se por desenvolver cada um destes modelos e analisar os resultados alcançados por cada um deles.

Após a determinação da função discriminante, será preciso aferir a adequação geral do modelo gerado. Em primeiro lugar, a matriz de classificação será construída, e será verificado o percentual de acerto de classificação para cada grupo e para o modelo como um todo. Em seguida, convém testar a hipótese de que o modelo desenvolvido é realmente um modelo cujo desempenho pode ser considerado melhor que uma classificação ao acaso, que é de 50%, uma vez que existem apenas 2 grupos em que os casos podem ser classificados.

Para comparar o índice de acerto alcançado com o índice de classificação ao acaso (50%), o teste estatístico Q de Press (Hair *et alii*, 1996, p.270). Esta medida compara o número de classificações corretas com o número total de casos na amostra

e o número de grupos. O valor de Q segue uma distribuição Qui-Quadrado com 1 grau de liberdade, e pode ser calculado a partir da seguinte expressão:

$$Q = \frac{[N - (nK)]^2}{N(K - 1)}, \text{ onde:}$$

N = tamanho total da amostra;

n = número de casos classificados corretamente e

K = número de grupos.

3.3.1.5 *Interpretação dos Resultados*

Após alcançados os resultados anteriores, é importante interpretar adequadamente os achados com o uso da técnica. O valor absoluto dos pesos padronizados dos coeficientes de discriminação fornecem a informação de quais variáveis possuem a maior capacidade contributiva em termos de poder de discriminação entre os dois grupos. Os sinais destes coeficientes indicam se a contribuição das variáveis para o modelo é positiva ou negativa. Uma outra forma de avaliar a importância individual das variáveis é através das cargas discriminantes, fornecida pela matriz de estrutura. Os valores destas cargas medem a correlação linear simples entre cada variável independente e a função discriminante canônica padronizada.

3.3.1.6 *Validação dos Resultados*

A última etapa do processo de determinação de uma função discriminante pelo método da ADL é a validação do modelo final. Como já mencionado na seção 3.3.1.2,

devido à quantidade de dados insuficientes para se formar uma sub-amostra do tipo *holdout*, será adotada a validação do tipo *leave-one-out*, onde $n - 1$ casos são utilizados para desenvolver o modelo discriminante a fim de classificar o caso n .

3.3.2 *A Regressão Logística*

Assim como a ADL, a RL é apropriada quando a variável dependente é do tipo não métrica. Devido aos pressupostos rígidos da ADL, a RL tornou-se preferida em estudos desta natureza. A RL se parece muito com uma regressão múltipla e por isso também seu uso é bastante apreciado entre os pesquisadores. Ela permite também a incorporação de efeitos não lineares, e há uma série de diagnósticos disponíveis para a análise dos resultados.

3.3.2.1 *Representação da Variável Dependente Binária*

Ao contrário da ADL, a RL não baseia suas predições em escores discriminantes. A RL aborda os mesmos tipos de problemas que a ADL, e de uma forma mais parecida com a regressão múltipla. A diferença é que a RL prediz diretamente a probabilidade de um evento ocorrer, que pode ser qualquer valor entre zero e um. Os valores preditos devem estar limitados ao intervalo de zero e um, e para definir essa relação, a RL utiliza uma relação entre a variável dependente e as variáveis independentes que se assemelha a uma curva em forma de um S, conforme pode ser visto na figura a seguir. Para valores muito baixos da variável independente,

a probabilidade se aproxima de zero. À medida que o valor da variável independente aumenta, a probabilidade aumenta rapidamente, mas devido à característica da curva, passa a aumentar lentamente e tende assintoticamente para o valor 1, mas nunca o ultrapassa.

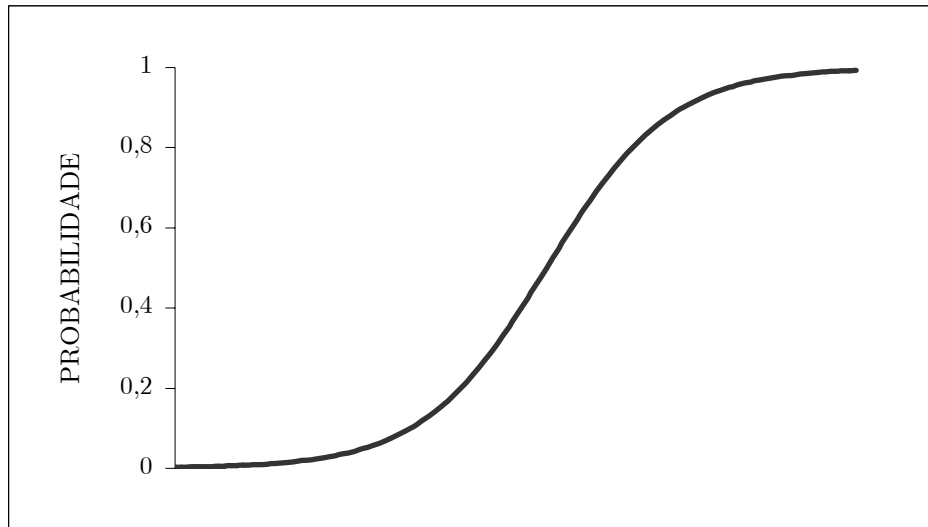


Figura 6: Exemplo de Regressão Logística.

3.3.2.2 Estimação do Modelo de Regressão Logística

Ao contrário da regressão múltipla a RL não usa o método dos mínimos quadrados na determinação dos coeficientes da regressão. A natureza não linear da transformação logística requer que um outro procedimento, o da máxima verossimilhança, seja usado de uma maneira iterativa para encontrar as estimativas mais prováveis para seus coeficientes.

O processo de estimação dos coeficientes é, no entanto muito semelhante ao da própria regressão linear. Para estimar o modelo de regressão logística, a curva em S

mostrada anteriormente é ajustada aos dados reais que são observados da amostra. A figura a seguir mostra duas curvas de RL. Na primeira os dados se ajustam bem ao modelo, enquanto que na segunda o ajustamento não é bom. Estes dados representam os eventos dicotômicos em que a empresa está em dificuldades financeiras ou não. Na figura à esquerda os dados se ajustam bem, pois não há sobreposição dos eventos para o mesmo nível da variável dependente em questão. Na figura à direita a sobreposição é bem clara, e conseqüentemente o ajuste é fraco.

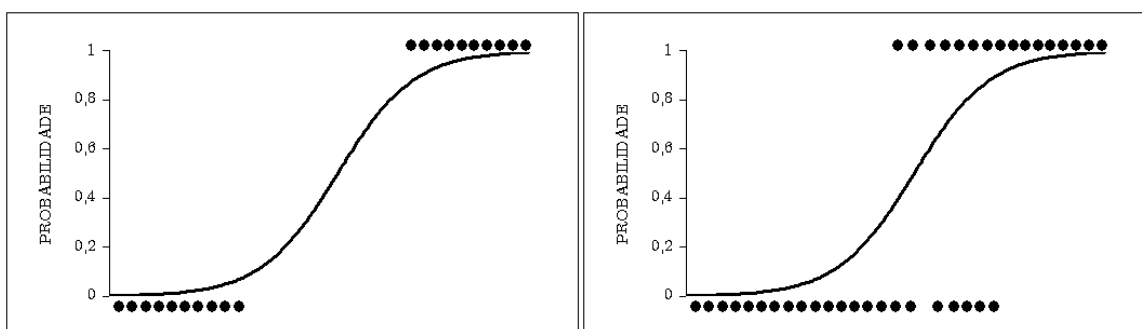


Figura 7: Curvas de Regressão Logística, respectivamente bem ajustada e mal ajustada.

O critério de seleção de variáveis que compõem o modelo final da regressão logística se assemelha ao descrito para o modelo discriminante: foi utilizado o procedimento de todas as regressões possíveis (*all-possible-regressions procedure*). Como dito anteriormente, este procedimento visa encontrar um ou mais subconjuntos de variáveis a fim de modelar o problema de pesquisa.

A fim de verificar a significância do modelo final, existem dois testes possíveis de serem conduzidos (Hair *et alii*, 1998, p.317). O primeiro deles baseia-se no desvio entre o modelo ajustado e o que se chama de modelo saturado, ou seja, um modelo onde n parâmetros ajustam perfeitamente n observações. Quanto menor for o desvio, medido pela estatística $-2LL$ ($-2 \loglikelihood$), melhor é o modelo. O outro teste

conduzido é o teste de Hosmer e Lemeshow. Este teste indica se há ou não diferença estatística entre as classificações esperadas e as observadas no modelo.

3.3.2.3 *Interpretação dos Coeficientes*

Uma das vantagens da RL é a de que é preciso saber apenas se um evento ocorreu para então usar o valor dicotômico como variável dependente. A partir deste valor dicotômico o método prediz sua estimativa para a probabilidade de que o evento venha a ocorrer ou não.

O procedimento para calcular os coeficientes da RL compara a probabilidade do evento ocorrer com a probabilidade do evento não ocorrer. Essa razão pode ser expressa através da seguinte equação:

$$\frac{\text{Prob}_{(\text{evento})}}{\text{Prob}_{(\text{sem evento})}} = e^{B_0 + B_1 X_1 + \dots + B_n X_n}$$

Os coeficientes $B_0, B_1, B_2, \dots, B_n$ são na verdade medidas das mudanças na razão de probabilidades apresentada acima. Eles são expressos em logaritmos de forma que eles podem ser transformados de volta (através do antilog) para que seu efeito relativo nas probabilidades seja percebido mais facilmente. Assim, um coeficiente positivo aumenta a probabilidade, enquanto que um coeficiente negativo diminui a probabilidade.

3.3.2.4 Validação dos Resultados

A fim de validar o modelo de RL desenvolvido, a amostra inicial de 40 casos foi separada de 3 formas diferentes em dois sub-grupos: um grupo de classificação e um grupo de validação. A separação se deu por meio de sorteio dos casos que compõem cada um dos grupos, e estas três formas foram chamadas de *Selec80*, *Selec 60* e *Selec 50*. A *Selec80* possui 80% dos casos na amostra de classificação e 20% na amostra de validação. A *Selec60* possui 60% dos casos na amostra de classificação e a *Selec50* possui a mesma quantidade de empresas em cada um dos grupos.

Para cada uma das formas apresentadas, foi desenvolvida uma equação logística onde os casos da amostra de validação foram testados. Os resultados da classificação para o grupo de empresas de validação conferem ao processo uma medida da capacidade de generalização da expressão obtida com o uso da técnica.

3.3.3 Redes Neurais Artificiais

As RNA são um sistema adaptativo capaz de aprender relações entre variáveis através dos dados a ela disponibilizados e realizar generalizações para outros dados ainda novos para a rede. Este processo é feito ao se disponibilizar uma série de dados de entrada com as respectivas saídas desejadas. A rede busca então aprender as relações entre os dados de entrada e de saída variando seus parâmetros. No caso da pesquisa desta dissertação, a rede a ser criada será utilizada com o objetivo de classificação de casos entre os dois grupos de empresas definidos. Será adotado como

saída da rede um valor entre zero e um, que representará a probabilidade de cada caso pertencer a um grupo ou outro.

A seleção de uma arquitetura de rede é muito importante. As RNA são basicamente divididas em dois grandes grupos: redes *feed-forward* e redes *feed-backward*. Uma rede neural se constitui de um conjunto de nós interligados, e a diferença básica entre esses dois métodos é na forma pela qual esses nós se interligam. As redes *feed-backward* possuem apenas uma camada de neurônios e todos são interligados entre si. As redes *feed-forward* possuem várias camadas organizadas horizontalmente, e cada neurônio conecta-se e envia informação para os demais da camada posterior. Os neurônios de uma mesma camada não são interligados entre si.

As redes *feed-forward* são freqüentemente compostas de três camadas: uma de entrada, uma intermediária e uma de saída. No entanto, mais camadas intermediárias podem ser definidas dependendo da complexidade dos dados do problema em questão. Os neurônios da camada de entrada introduzem a informação na rede e os de saída transmitem as respostas.

Nesta dissertação será usado o método de redes do tipo *feed-forward* com três camadas, baseada no experimento de Odom e Sharda (1993) descrito na seção 2.4.4.1. A figura a seguir apresenta a estrutura da rede a ser usada nesta pesquisa.

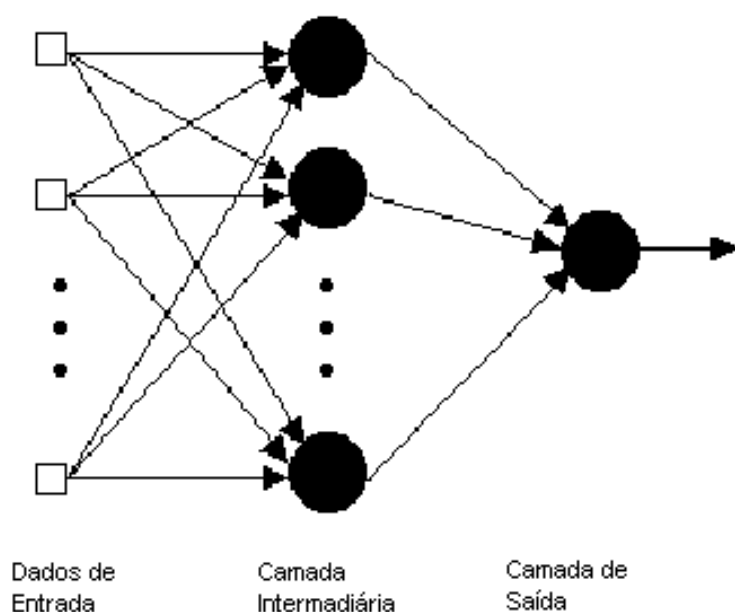


Figura 8: Arquitetura da rede a ser utilizada neste estudo.

Em uma rede como a mostrada na figura acima, há duas importantes características. Em primeiro lugar, os elementos de processamento não são lineares, e são definidos por funções de transformação que devem ser suaves, como uma função logística, por exemplo. Em segundo lugar, eles são intensamente interconectados, de forma que todos os elementos de uma camada se conectam a todos os elementos da camada seguinte.

No processo de criação de uma RNA a primeira etapa é a de aprendizado da rede. Nesta etapa, os casos a serem usados como dados de treinamento serão apresentados à rede por um determinado número de ciclos (*epochs*), e em cada ciclo de apresentação dos dados, a rede ajusta seus parâmetros de forma que consiga classificar corretamente o maior número de casos possíveis.

Para começar o processo de retro-propagação, um valor inicial será atribuído para cada peso de conexão e o processo de ajustamento continuará até que o critério de parada seja atendido. Há três critérios possíveis de serem adotados: número

máximo de iterações, erro quadrado médio¹⁹ do conjunto de dados de treinamento e o erro quadrado médio do conjunto de dados de validação²⁰. O método mais poderoso é sem dúvidas o baseado no erro quadrado médio do conjunto de dados de validação, pois ele interrompe o treinamento da rede no ponto de generalização ótima, minimizando o erro do conjunto de validação. O uso do conjunto de dados de validação também se justifica pela tentativa de evitar excesso de treinamento da rede, que faria com que ela memorizasse os dados de treinamento, mas que não fosse útil para aplicações em dados que não aqueles (generalização).

A função mais importante das RNA é a de aprender a partir dos dados disponíveis. Assim, os pesos de conexão entre cada elemento de processamento sofrem mudanças a partir de seus valores anteriores por meio de um termo de correção, que é especificado de acordo com a regra de aprendizagem. O foco da regra de aprendizagem é minimizar as diferenças entre os dados de saída obtidos através da passagem dos dados pela rede e os dados reais conhecidos *a priori*. O processo de aprendizagem leva em consideração a taxa de aprendizagem, que se for muito baixa faz com que o processo se torne demorado. Do contrário, se a taxa for muito alta, o processo de adaptação não converge adequadamente e os pesos passam a não ser calculados de forma útil à solução do problema.

¹⁹ O erro quadrado médio é obtido a partir da média dos quadrados das diferenças entre os dados de saída obtidos no processamento da rede e os dados de saída esperados.

²⁰ Neste caso, o termo validação se refere a uma porção dos dados não utilizados no conjunto de treinamento (normalmente cerca de 10% dos casos disponíveis), para servir como critério de parada. Além do conjunto de validação, há ainda o conjunto de teste, que serve como validação do modelo treinado originalmente. Não é conveniente utilizar os dados de teste como critério de parada, pois a rede criada será adequada para estes dados, e o intuito de uma validação seria perdido.

Outro aspecto importante a ser considerado na etapa de aprendizagem da rede é que não há garantias de que a solução final dos pesos das conexões será obtida. Uma RNA pode obter diferentes soluções para o mesmo conjunto de dados quando diferentes valores são usados inicialmente para os pesos de conexão. Assim, com vistas a se obter uma boa rede, ela deve ser treinada múltiplas vezes. O objetivo aqui é criar uma rede onde cada uma das rodadas de treinamento convirja para o mesmo erro quadrado médio.

Após a etapa de aprendizagem, a rede está pronta para a etapa de teste usando novos dados ainda desconhecidos pela rede. Usando os parâmetros obtidos na etapa anterior, a rede busca classificar novos dados em um dos grupos de interesse. Nesta etapa, os resultados alcançados pela rede são de vital importância, pois são eles que comprovam a capacidade de generalização da rede.

No processo de criação da rede, a definição de alguns parâmetros é importante para seu bom desempenho como uma ferramenta de generalização. Em primeiro lugar deve-se definir a quantidade de camadas intermediárias a serem usadas. Foi definida a utilização de uma camada intermediária, suficiente para fornecer um caráter não linear ao problema. O segundo passo é definir quantos elementos de processamento (ou neurônios) devem ser usados nesta camada. Para responder a esta pergunta, ainda durante a etapa de aprendizagem, foram treinadas diversas redes variando-se a quantidade de neurônios na camada intermediária a fim de encontrar o número de elementos de processamento mais adequado para o problema.

Foi realizada também uma análise de sensibilidade para identificar variáveis que acrescentaram pouca informação ao modelo. Esta etapa é semelhante a qualquer

outro método de seleção de variáveis em um modelo de regressão, como o método *stepwise* ou de melhores sub-conjuntos, e foi de grande importância na redução das variáveis iniciais do estudo.

3.4. Pacotes Estatísticos Utilizados

Para realizar a Análise Discriminante Linear e a Regressão Logística, foi utilizado o pacote estatístico *SPSS (Statistical Package for Social Sciences)* para Windows versão 9.0. Para o desenvolvimento das Redes Neurais Artificiais foi usado o pacote *Neuro Solutions* versão 4.20. Como suporte e também para a realização de alguns dos testes estatísticos como o de normalidade e para as transformações de Box-Cox foi utilizado o *Minitab* versão 13.0.

4. Resultados da Pesquisa

4.1. Análise Discriminante

4.1.1 *Análise de Normalidade dos Dados*

A fim de atender ao pressuposto de normalidade das variáveis independentes que fazem parte do estudo, inicialmente foram feitos quatro diferentes testes para verificar a normalidade ou não dos dados para cada variável: testes de Anderson-Darling (A-D), de Ryan-Joiner (R-J), de Shapiro-Wilk (S-W) e o de Kolmogorov-Smirnov (K-S) com a correção de Lilliefors, usando-se um nível de significância de 0,05. Em seguida, para as variáveis que não se apresentaram normalmente distribuídas, foi tentada uma transformação de Box-Cox, com o intuito de normalizar a distribuição de seus dados.

Os resultados dos testes de normalidade para as variáveis sem transformação estão listados na tabela 17 a seguir. Os resultados dos testes de A-D e de R-J (não mostrados) corroboram com os testes de S-W e de K-S, ou seja, apenas as variáveis X_{03} e X_{05} foram consideradas normalmente distribuídas. A única divergência entre os resultados dos testes de normalidade se deu com relação à variável X_{06} , que foi considerada normalmente distribuída pelo teste de K-S, mas não o foi pelos outros três testes.

Nas variáveis onde não foi possível comprovar a normalidade da distribuição de seus dados foi aplicado o procedimento de Box-Cox, que busca um fator λ ótimo a fim de tornar a distribuição da variável aproximadamente normal. As únicas variáveis

onde foi possível obter sucesso com a transformação de Box-Cox foram as variáveis X_{04} e X_{06} .

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	gl	p-valor	Estatística	gl	p-valor
X01	,382	40	,000	,383	40	,010**
X02	,382	40	,000	,452	40	,010**
X03	,099	40	,200*	,963	40	,348
X04	,167	40	,006	,879	40	,010**
X05	,086	40	,200*	,961	40	,316
X06	,135	40	,065	,914	40	,010**
X07	,501	40	,000	,188	40	,010**
X08	,470	40	,000	,393	40	,010**
X09	,511	40	,000	,172	40	,010**
X10	,398	40	,000	,434	40	,010**
X11	,394	40	,000	,356	40	,010**
X12	,482	40	,000	,182	40	,010**
X13	,456	40	,000	,203	40	,010**
X14	,322	40	,000	,547	40	,010**
X15	,407	40	,000	,397	40	,010**
X16	,424	40	,000	,221	40	,010**

** . Maior valor da significância real.

* . Menor valor da significância real.

a. Correção de Significância de Lilliefors.

Tabela 17: *Teste de Normalidade para as possíveis variáveis independentes.*

Os gráficos a seguir mostram os resultados do procedimento de Box-Cox para encontrar o valor de λ para as transformações das variáveis X_{04} e X_{06} . Os dados da última iteração para encontrar o valor de λ estão mostrados no canto superior direito do gráfico. O valor de λ estimado para a variável X_{04} foi de 0,225, significando que a variável deve ser elevada ao expoente λ . Para a variável X_{06} o

valor de λ encontrado foi de 0,337. A linha pontilhada vertical mostrada no gráfico limita o intervalo de confiança de 95% para o valor de λ . Isto significa que nos casos em que o valor estimado final é pouco interpretativo (como nos dois casos descritos), pode-se escolher um dos valores contidos neste intervalo com a garantia de que há 95% de chance de λ estar neste intervalo. Sendo assim, além do valor final estimado pelo procedimento, serão adotados ainda os valores de λ igual a 0,5 para a transformação de cada uma das variáveis. Uma transformação de variável com λ de 0,5 é o mesmo que obter a raiz quadrada dos dados para esta variável, tornando a interpretação desta nova variável mais fácil.

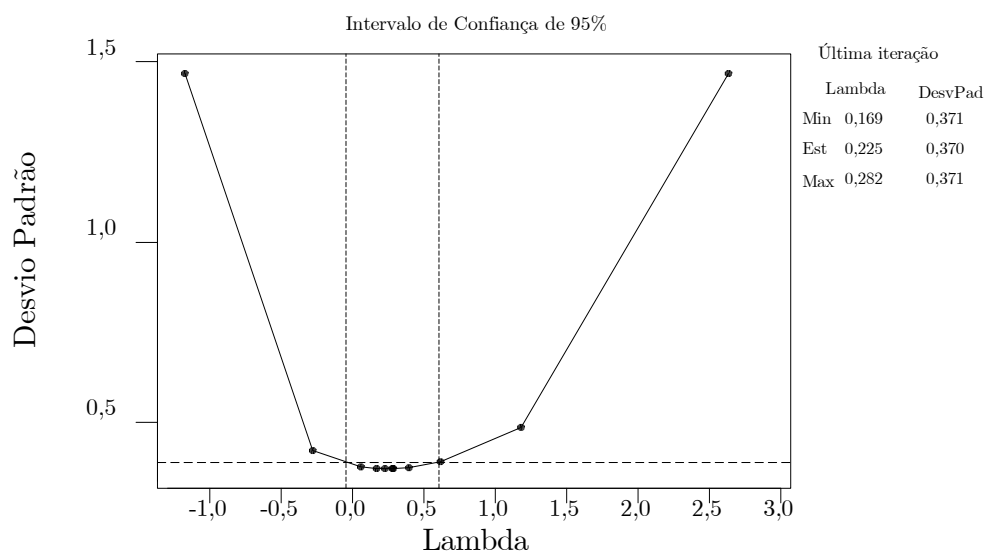


Gráfico 3: Transformação de Box-Cox para a variável X_{04} .

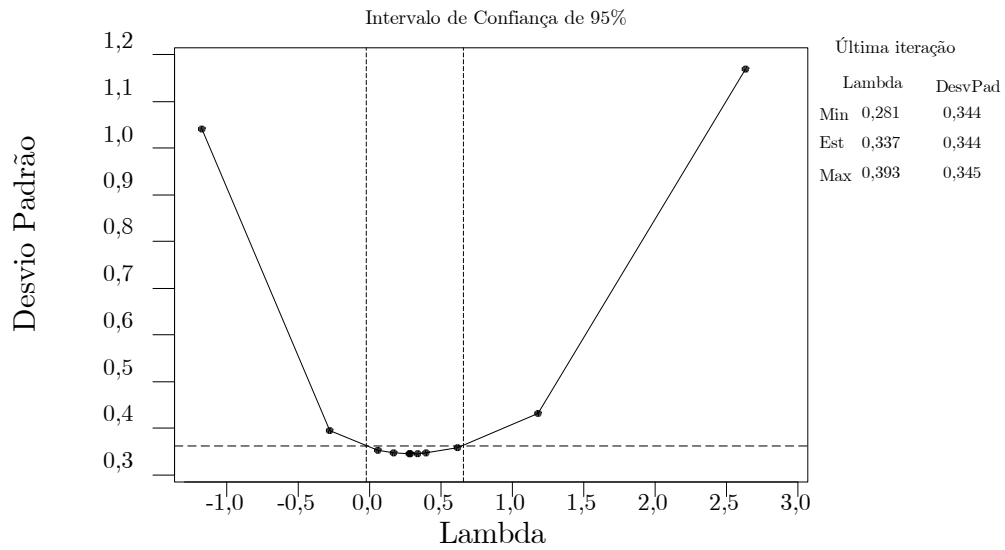


Gráfico 4: Transformação de Box-Cox para a variável X_{06} .

Após transformadas as variáveis, foi analisado novamente através dos quatro testes de normalidade anteriormente mencionados se a transformação empregada realmente resolveu o problema da falta de distribuição normal. Os dados da tabela abaixo apresentam os resultados para os testes de K-S e de S-W. Os testes de A-D e de R-J não estão apresentados, mas corroboram com os resultados dos outros dois testes. XT04 é a variável X_{04} transformada pelo λ estimado. A variável RAIZ_X04 é a variável X_{04} transformada pelo λ igual a 0,5. A mesma nomenclatura foi adotada para as transformações da variável X_{06} .

Para as quatro transformações realizadas, os quatro testes realizados asseguram a normalidade da distribuição dos dados.

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Estatística	gl	p-valor	Estatística	gl	p-valor
XT04	,088	40	,200*	,971	40	,478
RAIZ_X04	,099	40	,200*	,983	40	,831
XT06	,120	40	,152	,982	40	,820
RAIZ_X06	,092	40	,200*	,982	40	,805

*. Menor valor da significância real.

a. Correção de Significância de Lilliefors.

Tabela 18: Teste de Normalidade para as variáveis X_{04} e X_{06} transformadas.

4.1.2 Seleção das Variáveis do Modelo Discriminante

A fim de selecionar as variáveis que compõem o modelo discriminante linear com vistas à classificação dos casos do estudo entre empresas em situação de solvência e empresas em situação de insolvência, foi adotado o procedimento de redução de variáveis do tipo todas as regressões possíveis (*all-possible-regressions procedure*). Esse procedimento visa diminuir a quantidade de variáveis de um modelo a partir de todos os sub-conjuntos de variáveis possíveis de acordo com um critério específico. O procedimento busca escolher alguns poucos sub-conjuntos de variáveis que deverão ser então testados e cabe ao pesquisador escolher um ou mais que se adapte bem à solução do problema.

As variáveis usadas a partir de agora se resumem a apenas duas das dezesseis variáveis iniciais do estudo mais quatro transformações. Serão analisadas as variáveis X_{03} e X_{05} nas suas formas não transformadas e as variáveis X_{04} e X_{06} segundo as duas transformações sugeridas pelo procedimento de Box-Cox.

O resultado da análise para os possíveis sub-conjuntos de variáveis estão apresentados na tabela a seguir. O ‘X’ na coluna da variável significa que ela faz parte do modelo.

Vars	p	R ²	R ² (aj)	Cp	\sqrt{MSE}	X03	XT04	RAIZ_X04	X05	XT06	RAIZ_X06
1	2	25,6	23,7	1,4	0,44236					X	
1	2	24,8	22,8	1,8	0,44479						X
1	2	10,7	8,4	8,9	0,48472				X		
1	2	8,0	5,6	10,3	0,49211	X					
1	2	5,4	2,9	11,5	0,49889		X				
2	3	31,2	27,5	0,6	0,43125				X	X	
2	3	30,9	27,2	0,7	0,43207	X				X	
2	3	30,5	26,8	0,9	0,4333	X					X
2	3	30,4	26,6	1,0	0,43385				X		X
2	3	27,6	23,7	2,4	0,44231			X		X	
3	4	33,1	27,5	1,6	0,43103	X			X	X	
3	4	32,7	27,1	1,8	0,43245	X			X		X
3	4	31,3	25,6	2,5	0,43675			X	X	X	
3	4	31,2	25,5	2,6	0,43716		X		X	X	
3	4	31,2	25,5	2,6	0,43718				X	X	X
4	5	33,3	25,6	3,5	0,43663	X	X		X	X	
4	5	33,1	25,5	3,6	0,43711	X		X	X	X	
4	5	33,1	25,5	3,6	0,43714	X			X	X	X
4	5	33,0	25,4	3,7	0,43738	X	X		X		X
4	5	32,7	25,1	3,8	0,43835	X		X	X		X
5	6	34,3	24,6	5,0	0,43967	X	X	X	X		X
5	6	34,2	24,6	5,1	0,43983	X	X	X	X	X	
5	6	33,3	23,5	5,5	0,44293	X	X		X	X	X
5	6	33,1	23,3	5,6	0,44348	X		X	X	X	X
5	6	32,5	22,6	5,9	0,4455	X	X	X		X	X
6	7	34,3	22,4	7,0	0,44604	X	X	X	X	X	X

Tabela 19: Valores de R², R² ajustado, Cp e S para os sub-conjuntos do modelo.

Cada um dos quatro critérios constantes da tabela (R², R² ajustado, Cp e S) possui uma metodologia de análise individual, que será analisada em detalhe nas sub-seções a seguir.

4.1.2.1 Critério R^2

Segundo Neter *et alii* (1996), o critério do R^2 usa o coeficiente de determinação múltipla na busca do melhor sub-conjunto de variáveis. A regra inicial é a de que modelos cujos R^2 sejam altos são melhores do que quando são baixos. Como o valor de R^2 sempre cresce na medida em que novas variáveis são adicionadas ao modelo, deve-se analisar o ponto em que adicionar novas variáveis não representa um ganho substancial no valor de R^2 .

O gráfico abaixo ajuda na identificação do sub-conjunto para o qual o valor de R^2 apresenta um valor que não cresce substancialmente à medida que novas variáveis são adicionadas ao modelo. O gráfico deixa claro que pouco se ganha quando a quantidade de variáveis passa de quatro parâmetros. Da tabela 19 se verifica que o sub-conjunto mais adequado segundo este critério é o formado pelas variáveis X_{03} , X_{05} e X_{106} (marcado no gráfico a seguir com uma “o”).

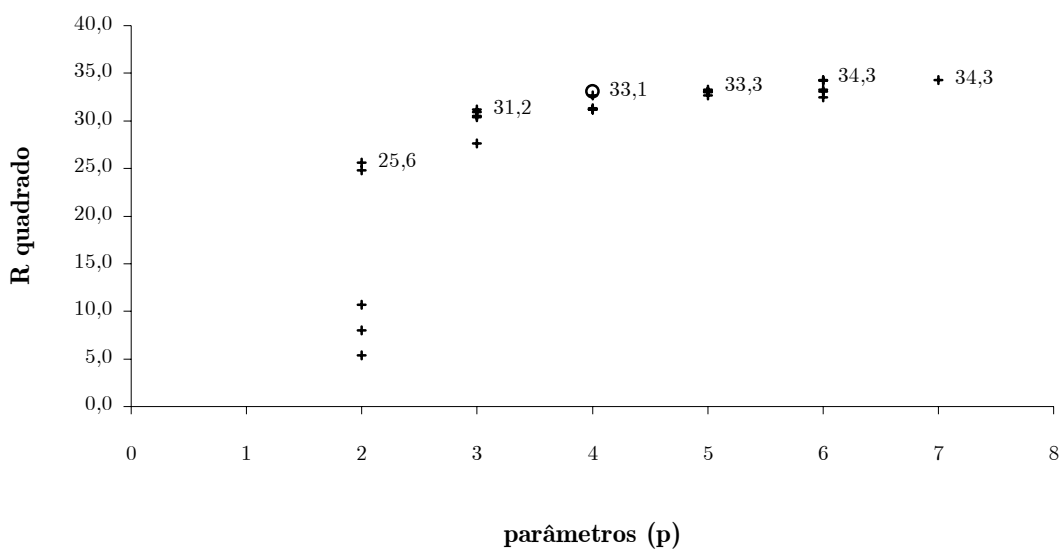


Gráfico 5: Gráfico de R^2 contra o número de parâmetros.

4.1.2.2 Critério do R^2 ajustado

Ao contrário do critério anterior, este agora leva em consideração a quantidade total de parâmetros no modelo. O critério de decisão se baseia no maior valor do R^2 ajustado ou tão próximo do máximo que não convém adicionar novas variáveis. O gráfico a seguir fornece uma ajuda adicional no processo de escolha.

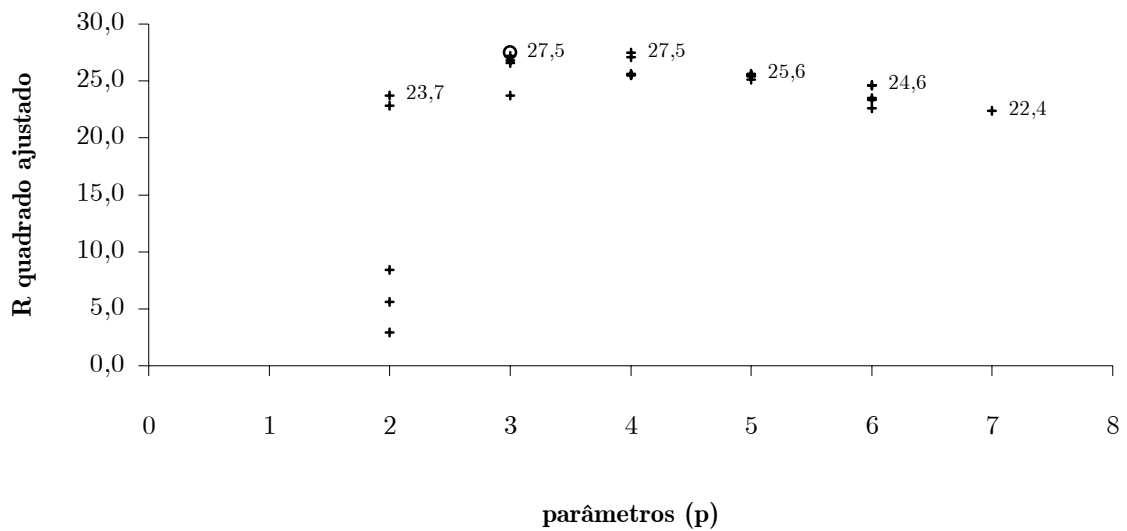


Gráfico 6: Gráfico de R^2 ajustado contra o número de parâmetros.

Da análise do gráfico e da tabela de dados, o melhor sub-conjunto de variáveis segundo este critério é o formado pelas variáveis X_{05} e X_{t06} . Convém notar que o valor do R^2 ajustado para este sub-conjunto é o mesmo valor para o sub-conjunto escolhido pelo critério anterior, o que significa que há semelhança nos resultados obtidos até o momento.

4.1.2.3 Critério C_p

Este terceiro critério está relacionado com o erro quadrado médio total, que é formado por dois componentes: um de natureza aleatória e outro de viés. O valor de C_p é o estimador da variância do modelo contendo $p - 1$ variáveis. Quando o valor de C_p é plotado em um gráfico contra a quantidade de parâmetros, os modelos que possuem menos viés são aqueles que tendem a ficar próximos da linha $C_p = p$. Modelos que possuem muito viés ficarão bastante acima desta linha, e os modelos que ficam abaixo desta linha são considerados sem viés, mas estão abaixo da linha por conta de erro amostral.

De acordo com Neter *et alii* (1996), usa-se o critério C_p na busca de identificar um sub-conjunto de variáveis que possuam um baixo valor de C_p e cujo valor seja próximo de p . Os sub-conjuntos com baixo valor de C_p possuem um erro quadrado médio total menor, e quando o seu valor é próximo a p quer dizer que o modelo possui menos viés.

Do gráfico abaixo é possível verificar que há vários sub-conjuntos abaixo da linha $C_p = p$. O sub-conjunto que possui a menor diferença entre o valor de C_p e de p é aquele formado pelas variáveis X_{03} , X_{t04} , raiz de X_{04} , X_{t06} e raiz de X_{06} . Apesar do resultado do critério apontar para este sub-conjunto, é possível dizer de antemão que este não é um conjunto de variáveis adequado, pois apresentará sérios problemas de multicolinearidade, uma vez que há duas variáveis que são combinações lineares de outras variáveis.

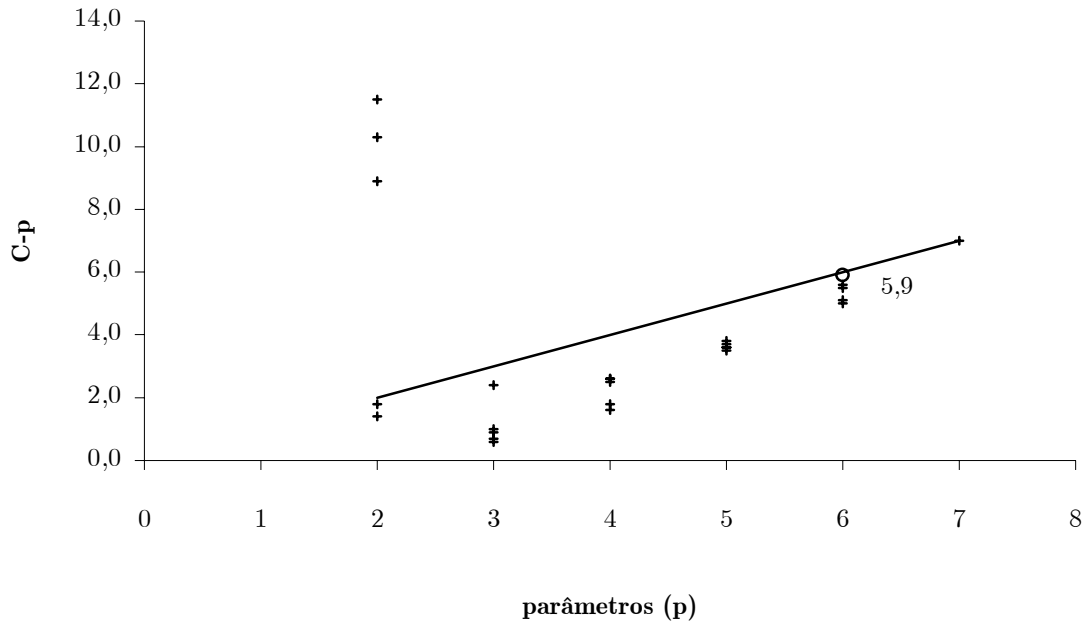


Gráfico 7: Gráfico de C_p contra o número de parâmetros.

4.1.2.4 Critério \sqrt{MSE}

O último critério a ser adotado é o do erro quadrado médio (MSE). O objetivo do critério é encontrar o sub-conjunto de variáveis com menor valor de \sqrt{MSE} ou aquele em que o acréscimo de variáveis não traz aumentos substanciais em seu valor. Sendo assim, de acordo com o gráfico a seguir, os sub-conjuntos das variáveis $[X_{03}, X_{05}$ e $X_{106}]$ e $[X_{05}$ e $X_{106}]$ foram os escolhidos. Como estes dois sub-conjuntos já foram escolhidos pelos critérios R^2 e R^2 ajustado respectivamente, nenhuma contribuição adicional pode ser obtida por este critério.

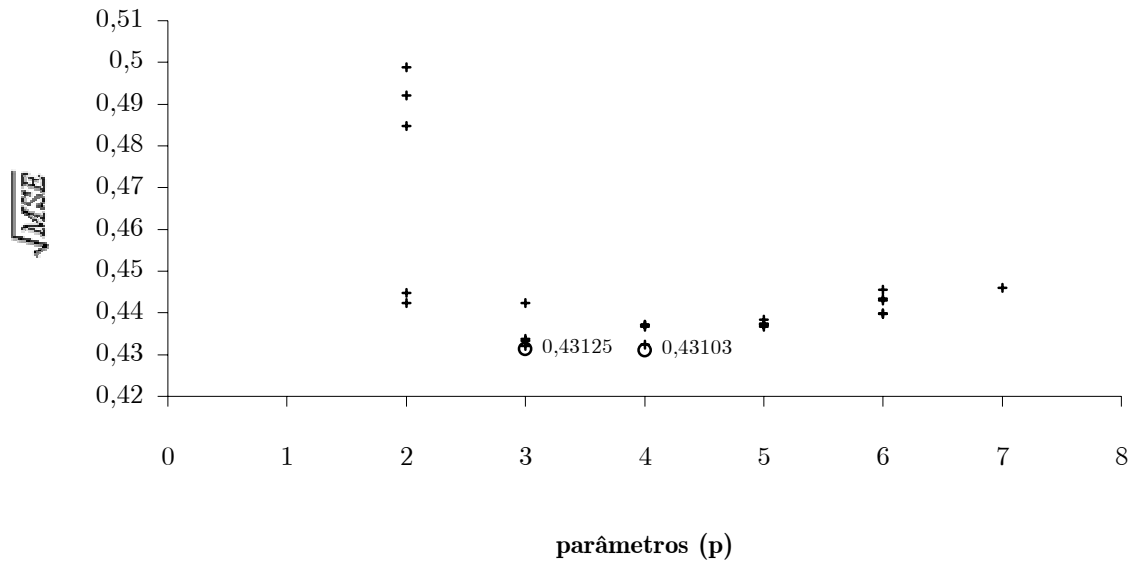


Gráfico 8: Gráfico de \sqrt{MSE} contra o número de parâmetros.

4.1.3 Análise de Multicolinearidade

Após a escolha dos dois sub-conjuntos de variáveis a serem testados na ADL, é preciso saber se há problemas de multicolinearidade entre as variáveis. A primeira forma de analisar o problema emprega o critério do Fator de Inflação de Variância (*VIF*). Valores de *VIF* maiores que 10 são um forte indicador de que há multicolinearidade entre as variáveis. A segunda forma é usando o Índice de Condição: valores maiores do que 15 indicam que pode haver problemas moderados de multicolinearidade entre variáveis. Valores maiores que 30 indicam que deve haver graves problemas de multicolinearidade.

Para ter certeza se há ou não problemas de multicolinearidade após o Índice de Condição ficar acima de 15, deve-se olhar para o coeficiente de proporção de

variância. Quando duas ou mais variáveis apresentarem valores maiores que 0,90, há multicolinearidade.

4.1.3.1 Sub-conjunto $[X_{03}, X_{05}, X_{106}]$

De acordo com os dados da Tabela 20 a seguir, não há problemas de multicolinearidade entre as variáveis. De acordo com a tabela 21, há uma dimensão em que o Índice de Condição ultrapassa a regra prática de 15, e portanto pode ser um indicio de que há multicolinearidade entre variáveis. Como apenas a variável X_{106} apresenta coeficiente de proporção de variância acima de 0,90, não se pode dizer que há problema de multicolinearidade para este sub-conjunto.

		Estatísticas de Colinearidade	
Modelo		Tolerância	VIF
1	X03	,789	1,268
	X05	,188	5,312
	XT06	,189	5,289

Tabela 20: *Estatísticas de Colinearidade.*

Modelo	Dimensões	Eigenvalue	Índice de Condição	Proporções de Variância			
				(Constante)	X03	X05	XT06
1	1	3,656	1,000	,00	,01	,00	,00
	2	,230	3,988	,00	,37	,08	,00
	3	,106	5,862	,10	,32	,13	,01
	4	8,106E-03	21,237	,89	,30	,79	,98

Tabela 21: *Eigenvalues, índice de condição e proporções de variância.*

4.1.3.2 Sub-conjunto $[X_{05}, X_{106}]$

Para o segundo sub-conjunto selecionado para a ADL, o diagnóstico de multicolinearidade feito através do *VIF* indica que não há problemas entre as variáveis (Tabela 22). Da análise dos Índices de Condição (Tabela 23) é possível identificar que há uma dimensão em que seu valor ultrapassa o valor de 15. No entanto, não há problemas de multicolinearidade pois apenas a variável X_{106} apresenta proporção de variância acima de 0,90.

Estatística de Colinearidade			
Modelo		Tolerância	VIF
1	X05	,236	4,238
	XT06	,236	4,238

Tabela 22: Estatísticas de Colinearidade.

Modelo	Dimensão	Eigenvalue	Índice de Condição	Proporção de Variância		
				(Constante)	X05	XT06
1	1	2,847	1,000	,01	,01	,00
	2	,141	4,489	,15	,23	,00
	3	1,133E-02	15,851	,85	,76	1,00

Tabela 23: Eigenvalues, índice de condição e proporções de variância.

4.1.4 *Desenvolvimento do Modelo Discriminante*

Uma vez aprovados os dois sub-conjuntos de variáveis, o próximo passo foi o de desenvolver o modelo discriminante. As seções a seguir apresentam este desenvolvimento para cada um dos dois sub-conjuntos analisados como úteis para o modelo discriminante final.

4.1.4.1 *ADL para o sub-conjunto $[X_{03}, X_{05}, X_{106}]$*

Inicialmente, as estatísticas descritivas para o conjunto de 40 empresas que fazem parte da amostra são mostradas na tabela a seguir. A Tabela 24 a seguir apresenta as estatísticas descritivas para as três variáveis selecionadas. Para a variável X_{03} , composição do endividamento, é possível notar que as empresas do grupo de insolventes (BKT) possuem uma maior quantidade de dívidas se vencendo no curto prazo. Para a variável X_{05} , índice de liquidez corrente, percebe-se que as empresas do grupo de solventes (NBK) são mais líquidas que as insolventes. Para a variável transformada X_{106} , índice de liquidez seca, é possível também notar que as empresas solventes são mais líquidas.

A Tabela 25 em seguida apresenta os valores de lamda de Wilks e os resultados da análise de variância (ANOVA) univariada usada para testar a diferença de média entre os grupos para as variáveis independentes. A variável X_{03} se mostrou significativa para a ANOVA ao nível de 10%, a variável X_{05} se mostrou significativa ao nível de 5% e a variável X_{106} se mostrou significativa ao nível de 1%.

	Status	X03	X05	XT06
Média	BKT	,5428	,6817	,6813
	NBK	,4085	1,0123	,9081
	Total	,4757	,8470	,7947
Desvio Padrão	BKT	,2925	,4915	,2220
	NBK	,1584	,4865	,1720
	Total	,2419	,5109	,2272

Tabela 24: Estatística descritiva das variáveis do modelo.

	Lambda de Wilks	F	gl1	gl2	p-valor
X03	,921	3,262	1	38	,079
X05	,893	4,571	1	38	,039
XT06	,744	13,051	1	38	,001

Tabela 25: Teste de igualdade de média para as variáveis do modelo.

Outro critério importante a ser atendido para que o modelo discriminante seja considerado adequado é a igualdade das matrizes de variância-covariância. Para testar esta igualdade é usado o teste M de Box, que para este sub-conjunto fornece o seguinte resultado, expresso na tabela 26 a seguir.

M de Box		15,033
F	Approx.	2,290
	gl1	6
	gl2	10462,189
	p-valor	,033

Tabela 26: Teste M de Box para igualdade das matrizes de variância-covariância.

Conforme pode ser visto pelos dados acima, as matrizes de variância-covariância não são diferentes ao nível de 5%, mas seriam ao nível de 1%.

As duas tabelas a seguir apresentam os aspectos multivariados do modelo discriminante. A Tabela 28 mostra que o modelo é altamente significativo (p-valor = 0,002). Na Tabela 27 é apresentada a correlação canônica do modelo, que é de 0,574, ou seja, 32,9% ($0,574^2$) da variância da variável dependente pode ser explicada por este modelo, que inclui apenas 3 variáveis independentes.

Função	Eigenvalue	% de Variância da Função	% de Variância Cumulativa	Correlação Canônica
1	,492	100,0	100,0	,574

Tabela 27: Função discriminante canônica.

Função	Lambda de			
	Wilks	Qui-Quadrado	gl	p-valor
1	,670	14,615	3	,002

Tabela 28: Significância da função discriminante canônica.

A Tabela 29 abaixo fornece os coeficientes de discriminação das funções canônicas. Os coeficientes não padronizados são aqueles usados para calcular o escore discriminante usado para classificação das empresas em um dos dois grupos.

	Coeficientes	
	padronizados	não padronizados
X03	-0,317	-1,346
X05	-0,686	-1,403
XT06	1,445	7,279
(Constante)		-3,956

Tabela 29: Coeficientes da função discriminante canônica.

As cargas da função discriminante são apresentadas na Tabela 30 a seguir. Nesta tabela, as variáveis estão ordenadas da carga mais alta até a carga mais baixa.

Os sinais das cargas discriminantes são coerentes com o esperado. Para a variável X_{03} (composição do endividamento), quanto menor a quantidade de dívida de curto prazo, maior a probabilidade de ser classificada como solvente. Para as outras duas variáveis de liquidez, quanto maior a liquidez, maior a probabilidade da empresa ser classificada como solvente.

A variável X_{t06} possui a maior carga discriminante, o que significa que esta variável tem o maior poder discriminatório entre as variáveis do estudo. Em seguida vem a variável X_{05} e só depois vem a variável X_{03} , que apresenta o menor poder de discriminação entre as três do modelo.

Cargas Discriminantes	
Função 1	
XT06	,835
X05	,494
X03	-,418

Tabela 30: Cargas da função discriminante.

A Tabela 31 apresenta as duas funções discriminantes de Fisher. Essas funções são usadas para classificação.

	Status	
	BKT	NBK
X03	17,672	15,831
X05	-20,251	-22,171
XT06	59,707	69,664
(Constante)	-18,925	-24,336

Tabela 31: Coeficientes das funções discriminantes lineares de Fisher.

A Tabela 32 apresenta os centróides dos dois grupos de empresas. O centróide representa a média dos escores das funções discriminantes para cada grupo. Nos casos em que a quantidade de observações para cada grupo é igual, os centróides apresentam valores absolutos iguais.

Centróide do Grupo	
Status	Função 1
BKT	-,684
NBK	,684

Tabela 32: *Centróides das funções discriminantes canônicas.*

Uma vez que os dois grupos usados no estudo possuem a mesma quantidade de casos (n=20), o escore discriminante é igual a zero. Assim, uma empresa é classificada como pertencente ao grupo de empresas solventes (NBK) quando possui escore discriminante maior que zero, e classificada como empresa insolvente (BKT) quando seu escore discriminante é menor que zero.

A capacidade de predição do modelo pode ser verificada através dos dados da Tabela 33. Para a amostra original, 70% das empresas insolventes foram corretamente classificadas. O percentual de acerto para as empresas solventes foi um pouco maior, de 80%. O índice de acerto global foi de 75% das empresas classificadas corretamente.

Para a amostra de validação, o índice de acerto para o grupo de empresas insolventes caiu para 60%. O índice de acerto para o grupo de empresas solventes permaneceu em 80%. O índice global de acerto na amostra de validação do tipo *leave-one-out* foi de 70%.

		Grupo Predito			
		Status	BKT	NBK	Total
Grupo Original	Contagem	BKT	14	6	20
		NBK	4	16	20
	%	BKT	70,0	30,0	100,0
		NBK	20,0	80,0	100,0
Validação	Contagem	BKT	12	8	20
		NBK	4	16	20
	%	BKT	60,0	40,0	100,0
		NBK	20,0	80,0	100,0

Tabela 33: Matriz de classificação final para o grupo original e para a validação.

Antes de analisar os casos incorretamente classificados pelo modelo desenvolvido, foi medida a capacidade de previsão do modelo em relação ao acaso a fim de se saber se o modelo desenvolvido classifica significativamente melhor do que uma classificação realizada ao acaso.

A chance de se classificar as empresas em um dos dois grupos no caso em que eles são do mesmo tamanho ($n=20$) é de 50%. A fim de comparar o índice de acerto obtido com o índice de acerto ao acaso, usou-se a estatística Q de Press em um teste que compara o número de classificações corretas com o tamanho total da amostra e o número de grupos. O valor de Q segue uma distribuição qui-quadrado com 1 grau de liberdade. O valor de Q é calculado através da expressão:

$$Q = \frac{[N - (nK)]^2}{N(K-1)}, \text{ onde}$$

N = tamanho total da amostra

n = número de observações classificadas corretamente e

K = número de grupos.

Para os resultados obtidos com este modelo, tem-se $Q = 10$ (p-valor = 0,0016), sendo, portanto, bastante significativo para a amostra original. Para os resultados da amostra de validação, o valor de Q é de 6,4 (p-valor = 0,011), também bastante significativo.

A seguir, as Tabelas 34 e 35 apresentam os diagnósticos de classificação dos casos para a amostra original e para a amostra de validação. Da amostra original 10 casos não foram classificados corretamente. Os casos 3, 10, 12 e 15 deveriam ter sido classificados como solventes, mas foram classificados como insolventes. Os casos 26, 29, 33, 36, 37 e 38 deveriam ter sido classificados como insolventes, mas na verdade foram classificados como solventes.

Para a amostra de validação, além dos casos já descritos como classificados incorretamente, os casos 34 e 35 também foram classificados como solventes quando na verdade deveriam ter sido classificados como insolventes.

A fim de aprofundar a busca pelas razões das classificações incorretas, os casos classificados incorretamente foram combinados entre si e comparados aos casos classificados corretamente na tentativa de identificar diferenças específicas nas variáveis independentes. A Tabela 36 mostra que os 8 casos pertencentes ao grupo de empresas insolventes classificados incorretamente seja na amostra original, seja na amostra de validação, apresentam diferenças significativas para a variável X_{t06} .

Dados de Classificação						
	Caso	Grupo Real	Grupo Predito	Probabilidade (Grupo Predito = Grupo Real)	Probabilidade (<> Grupo Real)	Escore Discriminante
Original	1	1	1	,781	,219	,929
	2	1	1	,914	,086	1,726
	3	1	0	,516	,484	-,046
	4	1	1	,753	,247	,816
	5	1	1	,898	,102	1,588
	6	1	1	,868	,132	1,376
	7	1	1	,924	,076	1,830
	8	1	1	,571	,429	,209
	9	1	1	,666	,334	,504
	10	1	0	,587	,413	-,258
	11	1	1	,835	,165	1,185
	12	1	0	,601	,399	-,301
	13	1	1	,671	,329	,521
	14	1	1	,707	,293	,643
	15	1	0	,508	,492	-,023
	16	1	1	,859	,141	1,324
	17	1	1	,648	,352	,446
	18	1	1	,655	,345	,470
	19	1	1	,634	,366	,400
	20	1	1	,615	,385	,343
	21	0	0	,925	,075	-1,841
	22	0	0	,934	,066	-1,934
	23	0	0	,714	,286	-,669
	24	0	0	,875	,125	-1,420
	25	0	0	,868	,132	-1,376
	26	0	1	,733	,267	,739
	27	0	0	,913	,087	-1,721
	28	0	0	,971	,029	-2,569
	29	0	1	,804	,196	1,032
	30	0	0	,984	,016	-3,008
	31	0	0	,881	,119	-1,465
	32	0	0	,868	,132	-1,376
	33	0	1	,867	,133	1,374
	34	0	0	,510	,490	-,030
	35	0	0	,518	,482	-,053
	36	0	1	,607	,393	,319
	37	0	1	,547	,453	,138
	38	0	1	,821	,179	1,115
	39	0	0	,635	,365	-,404
	40	0	0	,674	,326	-,530

Tabela 34: Predições individuais dos casos para a amostra original.

Dados de Classificação					
	Caso	Grupo Real	Grupo Predito	Probabilidade (Grupo Predito = Grupo Real)	Probabilidade (Grupo Predito <> Grupo Real)
Validação	1	1	1	,769	,231
	2	1	1	,910	,090
	3	1	0	,552	,448
	4	1	1	,734	,266
	5	1	1	,893	,107
	6	1	1	,854	,146
	7	1	1	,917	,083
	8	1	1	,563	,437
	9	1	1	,648	,352
	10	1	0	,623	,377
	11	1	1	,814	,186
	12	1	0	,681	,319
	13	1	1	,658	,342
	14	1	1	,701	,299
	15	1	0	,635	,365
	16	1	1	,850	,150
	17	1	1	,628	,372
	18	1	1	,624	,376
	19	1	1	,609	,391
	20	1	1	,600	,400
	21	0	0	,919	,081
	22	0	0	,930	,070
	23	0	0	,687	,313
	24	0	0	,839	,161
	25	0	0	,846	,154
	26	0	1	,781	,219
	27	0	0	,908	,092
	28	0	0	,975	,025
	29	0	1	,862	,138
	30	0	0	,988	,012
	31	0	0	,872	,128
	32	0	0	,859	,141
	33	0	1	,922	,078
	34	0	1	,564	,436
	35	0	1	,560	,440
	36	0	1	,653	,347
	37	0	1	,609	,391
	38	0	1	,863	,137
	39	0	0	,597	,403
	40	0	0	,592	,408

Tabela 35: Predições individuais dos casos para a amostra de validação.

Como era de se esperar, se há diferenças significativas para a variável X_{t06} , também haverá diferença significativa para a variável raiz de X_{06} , que não faz parte do modelo. Foi constatado ainda que as empresas classificadas incorretamente possuem uma liquidez geral (X_{t04}) mais alta que as empresas classificadas corretamente (significante ao nível de 5%). No entanto, para testar se esta era uma variável importante deixada de fora no modelo, constatou-se que não houve ganho algum quando esta foi inserida em um novo modelo de quatro variáveis. Os índices de classificação permaneceram os mesmos, a variável só se mostrou significante para o teste de diferença de médias entre os dois grupos ao nível de 15%, o teste M de Box teve sua significância reduzida a 0,019 e a carga discriminante da nova variável foi a mais baixa entre as quatro.

Teste t para Igualdade de Médias				
	Corretamente Classificados (N=12)	Incorretamente Classificados (N=8)	t	p-valor (bicaudal)
X03	,6352	,4042	1,835 ^b	,083
X05	,5301	,9092	-1,785 ^b	,091
XT04 ^a	,7713	,9360	-2,110	,050
RAIZ_X04	,6250	,8779	-1,668	,113
XT06	,5534	,8731	-4,448 ^b	,000
RAIZ_X06	,4270	,8259	-4,552	,000

a. Teste realizado para esta variável considerando que as variâncias não são iguais.

b. Variáveis incluídas na função discriminante.

Tabela 36: Diferença de média para as variáveis independentes entre empresas classificadas corretamente e incorretamente para o grupo de empresas insolventes.

A Tabela 37 a seguir mostra que os 4 casos pertencentes ao grupo de empresas solventes classificados incorretamente, seja na amostra original, seja na amostra de validação, também apresentaram diferenças significativas para a variável X_{t06} . Significância também foi verificada para a variável raiz de X_{06} . Nenhuma outra variável independente apresentou diferença estatisticamente significativa. As empresas classificadas incorretamente apresentaram índice de liquidez seca menor que as empresas solventes classificadas corretamente.

Teste t para Igualdade de Médias				
	Corretamente Classificados (N=16)	Incorretamente Classificados (N=4)	t	Sig. (2-tailed)
X03	,4181	,3700	,533 ^a	,601
X05	1,0864	,7162	1,394 ^a	,180
XT04	,9169	,9253	-,112	,912
RAIZ_X04	,8437	,8754	-,197	,846
XT06	,9531	,7284	2,695 ^a	,015
RAIZ_X06	,9381	,6342	2,507	,022

a. Variáveis incluídas na função discriminante.

Tabela 37: Diferença de média para as variáveis independentes entre empresas classificadas corretamente e incorretamente para o grupo de empresas solventes.

4.1.4.2 ADL para o sub-conjunto [X_{05} , X_{t06}]

Os dados descritivos para as variáveis integrantes deste modelo já foram apresentadas na Tabela 24 da seção anterior, e portanto não se faz necessário repetir os dados. Daquela tabela foi possível notar que as duas variáveis que agora fazem

parte do modelo são diferentes entre si ao nível de 5% (X_{05}) e 1% (X_{t06}). A variável X_{03} que só era significativa ao nível de 10% foi excluída.

Quanto à igualdade das matrizes de variância-covariância, este novo modelo apresenta-se significativo ao nível de 10% segundo o teste M de Box. Portanto, é possível afirmar que as matrizes são iguais e que a ADL é um método adequado para analisar o problema.

M de Box			6,481
F	Approx.		2,037
	gl1		3
	gl2	259920,000	
	p-valor		,107

Tabela 38: Teste M de Box para igualdade das matrizes de variância-covariância.

Quanto aos aspectos multivariados do modelo discriminante, a Tabela 40 a seguir mostra que o modelo é altamente significativo (p-valor = 0,001) e a Tabela 39 mostra que 31,1% ($0,558^2$) da variância da variável dependente pode ser explicada por este modelo, que agora inclui apenas duas variáveis independentes.

Função	Eigenvalue	% de Variância da Função	% de Variância Cumulativa	Correlação Canônica
1	,451	100,0	100,0	,558

Tabela 39: Função discriminante canônica.

Função	Lambda de Wilks	Qui-quadrado	gl	p-valor
1	,689	13,782	2	,001

Tabela 40: Significância da função discriminante canônica.

A Tabela 41 apresenta os coeficientes padronizados e não padronizados de discriminação das funções canônicas.

	Coeficientes	
	Padronizados	não padronizados
X05	-0,988	-2,021
XT06	1,731	8,719
(Constante)		-5,217

Tabela 41: Coeficientes da função discriminante canônica.

A Tabela 42 a seguir apresenta as cargas da função discriminante ordenadas da carga mais alta até a carga mais baixa. Os sinais das cargas discriminantes são coerentes com o esperado. Como as duas variáveis que agora estão no modelo são variáveis de liquidez das empresas, é de se esperar que quanto maior a liquidez, maior a probabilidade de a empresa ser classificada como solvente. A variável X_{106} possui a maior carga discriminante, o que significa que esta variável tem o maior poder discriminatório entre as variáveis do estudo.

Cargas Discriminantes	
Função 1	
XT06	,872
X05	,516

Tabela 42: Cargas da função discriminante.

As duas funções discriminantes de Fisher usadas na classificação das empresas são mostradas na tabela a seguir.

	Status	
	BKT	NBK
X05	-13,269	-15,916
XT06	45,676	57,095
(Constante)	-11,729	-18,562

Tabela 43: Coeficientes das funções discriminantes lineares de Fisher.

A Tabela 44 apresenta os centróides dos dois grupos de empresas, que representam as médias dos escores das funções discriminantes para cada grupo.

Status	Centróide do Grupo
	Função 1
BKT	-,655
NBK	,655

Tabela 44: Centróides das funções discriminantes canônicas.

Quanto à capacidade de predição do modelo, a tabela a seguir mostra que 65% das empresas insolventes foram corretamente classificadas. O percentual de acerto para as empresas solventes foi de 80%, o mesmo valor que no modelo anterior. O índice de acerto global caiu um pouco, ficando em 72,5% das empresas classificadas corretamente.

Para a amostra de validação, os valores encontrados pelo modelo anterior se mantiveram. O índice de acerto para o grupo de empresas insolventes foi de 60% e o índice de acerto para o grupo de empresas solventes foi de 80%. O índice global de acerto na amostra de validação foi de 70%.

		Grupo Predito			Total
		Status	BKT	NBK	
Grupo Original	Contagem	BKT	13	7	20
		NBK	4	16	20
	%	BKT	65,0	35,0	100,0
		NBK	20,0	80,0	100,0
Validação	Contagem	BKT	12	8	20
		NBK	4	16	20
	%	BKT	60,0	40,0	100,0
		NBK	20,0	80,0	100,0

Tabela 45: Matriz de classificação final para o grupo original e para a validação.

A medida da capacidade de previsão do modelo usando a estatística Q de Press apresentou os seguintes resultados:

$$Q = \frac{[N - (nK)]^2}{N(K-1)} = \frac{[40 - (29 \times 2)]^2}{40(2-1)} = 8,1$$

Para os resultados obtidos com este modelo, tem-se um p-valor de 0,0044, um valor bastante significativo para a amostra original. Para os resultados da amostra de validação, o valor de Q é de 6,4 (p-valor = 0,011), também bastante significativo.

As Tabelas 46 e 47 a seguir apresentam os diagnósticos de classificação dos casos para a amostra original e para a amostra de validação respectivamente. Da amostra original 11 casos não foram classificados corretamente. Os casos 3, 10, 12 e 15 deveriam ter sido classificados como solventes, mas foram classificados como insolventes. Os casos 26, 29, 33, 35, 37, 38 e 39 deveriam ter sido classificados como insolventes, mas na verdade foram classificados como solventes.

Dados de Classificação						
Caso	Grupo Real	Grupo Predito	Probabilidade (Grupo Predito = Grupo Real)	Probabilidade (Grupo Predito <> Grupo Real)	Escore Discriminante	
Original 1	1	1	,791	,209	1,017	
2	1	1	,898	,102	1,659	
3	1	0	,607	,393	-,332	
4	1	1	,760	,240	,880	
5	1	1	,876	,124	1,494	
6	1	1	,906	,094	1,732	
7	1	1	,890	,110	1,593	
8	1	1	,526	,474	,079	
9	1	1	,565	,435	,199	
10	1	0	,530	,470	-,093	
11	1	1	,861	,139	1,394	
12	1	0	,697	,303	-,638	
13	1	1	,720	,280	,720	
14	1	1	,668	,332	,534	
15	1	0	,563	,437	-,193	
16	1	1	,828	,172	1,198	
17	1	1	,676	,324	,563	
18	1	1	,712	,288	,691	
19	1	1	,520	,480	,060	
20	1	1	,669	,331	,538	
21	0	0	,912	,088	-1,783	
22	0	0	,944	,056	-2,159	
23	0	0	,795	,205	-1,036	
24	0	0	,869	,131	-1,442	
25	0	0	,724	,276	-,736	
26	0	1	,639	,361	,435	
27	0	0	,872	,128	-1,467	
28	0	0	,935	,065	-2,034	
29	0	1	,748	,252	,829	
30	0	0	,980	,020	-2,971	
31	0	0	,902	,098	-1,694	
32	0	0	,825	,175	-1,182	
33	0	1	,859	,141	1,380	
34	0	0	,629	,371	-,402	
35	0	1	,653	,347	,483	
36	0	0	,505	,495	-,016	
37	0	1	,681	,319	,579	
38	0	1	,822	,178	1,171	
39	0	1	,505	,495	,016	
40	0	0	,802	,198	-1,069	

Tabela 46: Predições individuais dos casos para a amostra original.

Dados de Classificação

	Caso	Grupo Real	Grupo Predito	Probabilidade (Grupo Predito = Grupo Real)	Probabilidade (Grupo Predito <> Grupo Real)
Validação	1	1	1	,781	,219
	2	1	1	,893	,107
	3	1	0	,628	,372
	4	1	1	,743	,257
	5	1	1	,870	,130
	6	1	1	,902	,098
	7	1	1	,875	,125
	8	1	1	,521	,479
	9	1	1	,561	,439
	10	1	0	,554	,446
	11	1	1	,848	,152
	12	1	0	,751	,249
	13	1	1	,715	,285
	14	1	1	,664	,336
	15	1	0	,685	,315
	16	1	1	,817	,183
	17	1	1	,661	,339
	18	1	1	,695	,305
	19	1	1	,514	,486
	20	1	1	,663	,337
	21	0	0	,903	,097
	22	0	0	,943	,057
	23	0	0	,789	,211
	24	0	0	,833	,167
	25	0	0	,719	,281
	26	0	1	,656	,344
	27	0	0	,864	,136
	28	0	0	,933	,067
	29	0	1	,796	,204
	30	0	0	,984	,016
	31	0	0	,897	,103
	32	0	0	,814	,186
	33	0	1	,914	,086
	34	0	0	,601	,399
	35	0	1	,670	,330
	36	0	1	,505	,495
	37	0	1	,702	,298
	38	0	1	,863	,137
	39	0	1	,514	,486
	40	0	0	,789	,211

Tabela 47: Predições individuais dos casos para a amostra de validação.

Para a amostra de validação, além dos casos já descritos como classificados incorretamente, o caso 36 também foi classificado como solvente quando na verdade deveria ter sido classificado como insolvente.

Comparando os casos classificados incorretamente por este modelo com o modelo anterior, pode-se perceber que há apenas uma divergência: o caso 39 aparece agora como classificado incorretamente, fato que não aconteceu quando foi usado o modelo anterior. No entanto, o caso 34 que na amostra de validação do modelo anterior foi apontado como classificado incorretamente, neste novo modelo não foi mal classificado nenhuma vez.

Os casos classificados incorretamente foram combinados entre si e comparados aos casos classificados corretamente na tentativa de identificar diferenças específicas nas variáveis independentes que possam ter causado a classificação incorreta. A tabela 48 a seguir mostra que os 8 casos pertencentes ao grupo de empresas insolventes classificados incorretamente, seja na amostra original, seja na amostra de validação, apresentam diferenças significativas para a variável X_{t06} que já está incluída no modelo e para a variável X_{05} que não faz parte deste modelo. A inclusão da variável X_{05} já foi testada no modelo anterior e foi possível observar que houve ganho de classificação na amostra original, quando uma empresa a mais foi classificada corretamente.

Teste t para Igualdade de Médias				
	Corretamente Classificados (N=12)	Incorretamente Classificados (N=8)	t	p-valor (bicaudal)
X03	,5823	,4835	,731	,474
X05	,4915	,9670	-2,361 ^b	,030
XT04 ^a	,7729	,9335	-2,046	,057
RAIZ_X04	,6281	,8734	-1,610	,125
XT06	,5405	,8924	-5,595 ^b	,000
RAIZ_X06	,4111	,8498	-5,747	,000

a. Teste realizado para esta variável considerando que as variâncias não são iguais.

b. Variáveis incluídas na função discriminante.

Tabela 48: Diferença de média para as variáveis independentes entre empresas classificadas corretamente e incorretamente para o grupo de empresas insolventes.

Como os quatro casos das empresas pertencentes ao grupo de solventes classificados incorretamente são os mesmos do modelo anterior, a tabela para a diferença de médias não foi apresentada nesta seção. As mesmas análises para o modelo anterior servem para este modelo.

4.2. Regressão Logística

A segunda técnica empregada na construção de modelos de classificação entre as empresas selecionadas para este estudo foi a Regressão Logística. Segundo Hair, Jr *et alii* (1998, p.246), esta técnica tem a vantagem de ser menos afetada do que a ADL quando seus pressupostos não são atendidos. Segundo os autores, isto é particularmente comum quanto à questão da normalidade da variável. Sendo assim,

para o desenvolvimento deste modelo, todas as variáveis foram empregadas como candidatas a serem selecionadas, inclusive as variáveis transformadas pela raiz da variável original usadas na Análise Discriminante. Decidiu-se por excluir do rol de candidatas as variáveis transformadas pela potência sugerida na transformação de Box-Cox por conta de sua dificuldade interpretativa.

O critério de seleção das variáveis empregado foi o mesmo da Análise Discriminante: o critério dos melhores sub-conjuntos. Segundo Hosmer e Lemeshow (2000, p.128), apesar de existir uma variação do algoritmo usado na regressão linear para seleção dos melhores sub-conjuntos em Regressão Logística, este não se faz necessário, e que é perfeitamente aceitável o uso do algoritmo desenvolvido para a regressão linear.

4.2.1 Seleção de Variáveis da Regressão Logística

Os dados da tabela abaixo apresentam os melhores sub-conjuntos de variáveis para cada quantidade de parâmetros possível dentro da amostra. Assim como foi feito na ADL, aqui também foram usados os critérios do R^2 , R^2 ajustado, C_p e \sqrt{MSE} . Cada um dos critérios foi analisado separadamente e estão apresentados nas seções a seguir, onde se encontram as justificativas para as escolhas feitas.

						Variáveis																		
Vars	p	R2	R2 aj	Cp	\sqrt{MSE}	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	R4	R6	
1	2	24,8	22,8	5,3	0,44479																			X
1	2	20,6	18,6	7,6	0,457						X													
2	3	30,5	26,8	4,2	0,4333			X																X
2	3	30,4	26,6	4,3	0,43385				X															X
3	4	37,4	32,1	2,4	0,4171			X											X					X
3	4	36,7	31,4	2,8	0,41944				X										X					X
4	5	41,3	34,6	2,3	0,40947			X				X								X				X
4	5	40,8	34	2,6	0,41136			X						X				X	X					X
5	6	43,2	34,9	3,2	0,40868			X				X					X	X		X				X
5	6	43,1	34,8	3,3	0,40899			X	X							X		X	X					X
6	7	47,9	38,5	2,6	0,39718	X	X	X					X							X				X
6	7	46,6	36,9	3,4	0,40237									X		X	X		X	X				X
6	7	46,5	36,7	3,4	0,40277					X	X	X		X					X					X
7	8	50	39	3,5	0,39533	X	X	X					X						X	X				X
7	8	49,9	39	3,5	0,39553	X	X	X			X	X							X					X
7	8	49,7	38,7	3,7	0,3966						X	X	X		X	X		X						X
8	9	52,4	40,1	4,2	0,3918						X	X	X		X	X		X	X					X
8	9	51,5	39	4,6	0,39539	X	X					X	X		X	X			X					X
8	9	51,5	39	4,7	0,39561	X	X	X			X	X						X	X					X
9	10	54,7	41,2	4,9	0,38844	X	X	X					X	X		X	X		X					X
9	10	54	40,2	5,3	0,39159				X	X	X	X	X		X	X		X		X		X		X
9	10	53,8	40	5,4	0,39228	X	X				X	X	X		X	X		X		X				X
10	11	56,8	41,9	5,7	0,38583	X	X				X	X	X	X		X	X		X					X
10	11	56,3	41,2	6	0,38831	X	X	X			X	X	X		X	X		X		X				X
10	11	56	40,8	6,2	0,38953			X	X	X	X	X			X	X		X		X		X		X
11	12	57,9	41,4	7,1	0,38754	X	X				X	X	X	X		X	X		X			X		X
11	12	57,7	41,1	7,2	0,38855	X	X	X			X	X	X	X		X	X		X					X
11	12	57,6	40,9	7,3	0,38922	X	X	X			X	X	X	X		X	X		X					X
12	13	59,9	42,1	8	0,38535	X	X				X	X	X	X		X	X	X		X	X			X
12	13	58,7	40,4	8,7	0,39097	X	X	X	X	X	X	X			X	X		X		X		X		X
12	13	58,7	40,3	8,7	0,39123	X	X	X	X	X	X	X			X	X	X		X		X			X
13	14	60,4	40,6	9,8	0,39019	X	X				X	X	X	X	X	X	X		X	X				X
13	14	60	40	10	0,39217	X	X				X	X	X	X	X	X	X		X	X				X
13	14	60	40	10	0,39239	X	X				X	X	X	X		X	X	X		X	X	X		X
14	15	60,7	38,6	11,6	0,39668	X	X	X	X	X	X	X	X	X	X	X	X		X		X			X
14	15	60,5	38,5	11,7	0,39726	X	X				X	X	X	X	X	X	X		X	X	X			X
14	15	60,5	38,4	11,7	0,39751	X	X	X	X	X	X	X	X	X	X	X			X	X	X			X
15	16	61,1	36,8	13,4	0,40259	X	X	X	X	X	X	X	X	X	X	X	X		X		X			X
15	16	61	36,6	13,4	0,40314	X	X	X	X	X	X	X	X		X	X	X		X	X	X			X
15	16	61	36,6	13,5	0,40332	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X			X
16	17	61,6	34,9	15,1	0,40845	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X			X
16	17	61,1	34,1	15,4	0,41101	X	X	X	X	X	X	X	X	X	X	X	X		X		X			X
16	17	61,1	34,1	15,4	0,41104	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X			X
17	18	61,8	32,3	17	0,4167	X	X	X	X	X	X	X	X	X	X	X	X	X		X	X	X		X
17	18	61,7	32,1	17,1	0,41739	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X		X
17	18	61,3	31,4	17,3	0,41931	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X			X
18	19	61,8	29,1	19	0,4265	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X

Tabela 49: Valores de R^2 , R^2 ajustado, C_p e S para os sub-conjuntos do modelo logístico.

4.2.1.1 Critério R^2

De acordo com este critério, deve-se escolher um sub-conjunto de variáveis quando o ganho incremental pelo aumento de uma nova variável ao modelo deixa de ser relevante. Analisando-se os ganhos existentes a cada acréscimo de variável ao modelo, foi escolhido o sub-conjunto formado por 6 variáveis $[X_1; X_2; X_3; X_8; X_{15}; X_{r6}]$.

4.2.1.2 Critério do R^2 ajustado

O uso do critério do R^2 ajustado permite encontrar aquele sub-conjunto de variáveis onde a tendência de crescimento se inverte, e o valor passa a diminuir. O gráfico a seguir mostra claramente esta inversão. Deve-se escolher como melhor sub-conjunto aquele que apresenta o maior valor da estatística ou aquele para o qual os ganhos com o aumento de novas variáveis ao modelo passam a não ser mais relevantes.

O maior valor encontrado dentre os possíveis sub-conjuntos para o R^2 ajustado foi de 42,1. Este sub-conjunto porém não foi escolhido para compor o grupo de sub-conjuntos selecionados. Escolheu-se o sub-conjunto das variáveis $[X_1; X_2; X_3; X_8; X_9; X_{11}; X_{12}; X_{15}; X_{r6}]$, que apresentou um R^2 ajustado igual a 41,2. Outros dois sub-conjuntos selecionados foram os formados pelas variáveis $[X_3; X_8; X_{15}; X_{r6}]$ e $[X_1; X_2; X_3; X_8; X_{15}; X_{r6}]$.

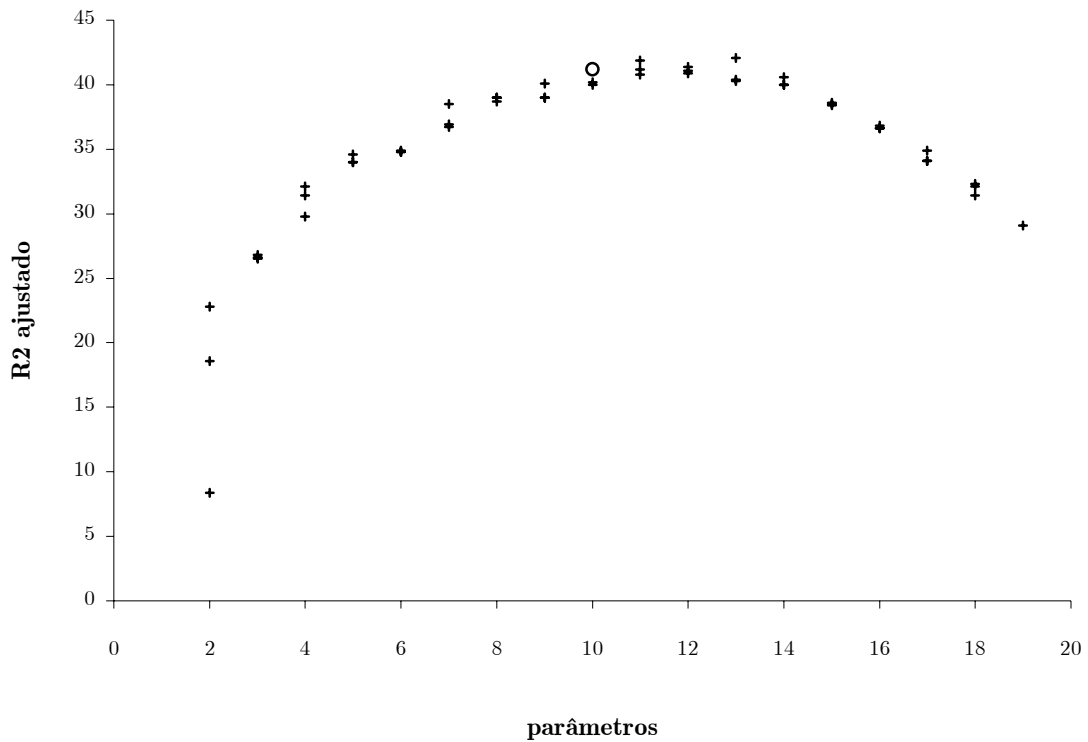


Gráfico 9: Gráfico de R^2 ajustado contra o número de parâmetros.

4.2.1.3 Critério C_p

Segundo o critério C_p , deve-se escolher o sub-conjunto de variáveis que mais se aproxime da linha onde C_p é igual ao número de parâmetros. Isto foi observado para o sub-conjunto formado por três variáveis $[X_3; X_{14}; X_{r6}]$. Este sub-conjunto está em detalhe no gráfico abaixo, apresentado com o marcador “o”. Outro sub-conjunto selecionado foi o sub-conjunto formado pelas variáveis $[X_3; X_8; X_{15}; X_{r6}]$, que apresenta o menor valor de C_p entre os sub-conjuntos listados na Tabela 49.

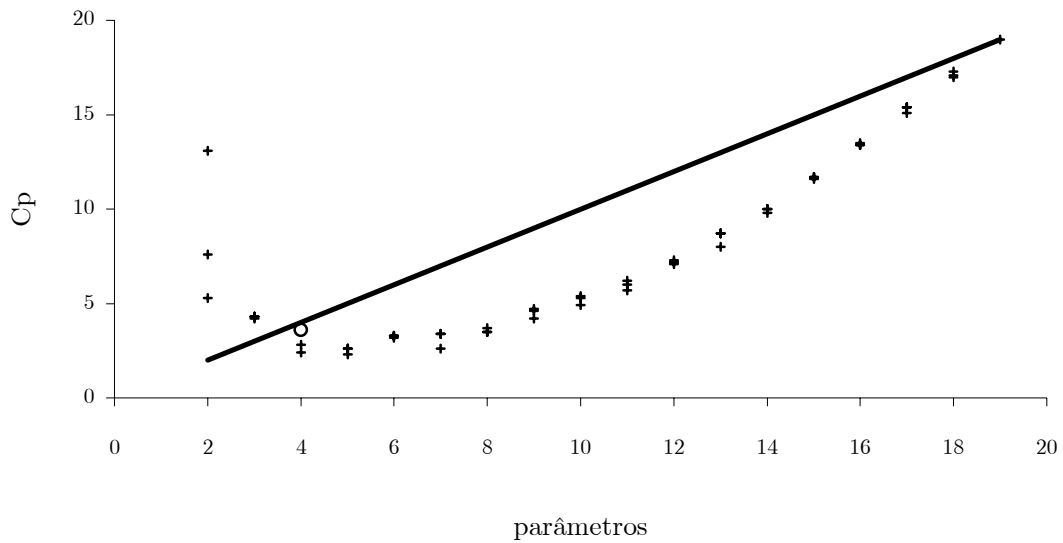


Gráfico 10: Gráfico de C_p contra o número de parâmetros.

4.2.1.4 Critério \sqrt{MSE}

Este critério diz que deve ser escolhido o sub-conjunto de variáveis que apresente o menor valor para \sqrt{MSE} ou aquele para o qual o acréscimo de novas variáveis não tragam mais reduções relevantes. Da análise dos dados da tabela 49, o menor valor de \sqrt{MSE} foi 0,38535. Porém, o sub-conjunto escolhido segundo o critério de relevância dos ganhos incrementais foi o composto pelas variáveis $[X_1; X_2; X_3; X_8; X_{15}; X_{r6}]$, mesmo sub-conjunto já escolhido segundo o critério do R^2 .

4.2.2 Análise de Multicolinearidade

A análise de multicolinearidade foi realizada para cada sub-conjunto pré-selecionado, utilizando-se dos Fatores de Inflação de Variância (*VIF*) e do Índice de Condição. As seções seguintes trazem as tabelas e as análises de multicolinearidade.

4.2.2.1 Sub-conjunto $[X_1; X_2; X_3; X_8; X_{15}; X_{r6}]$

Os dados da Tabela 50 a seguir indicam que há problemas de multicolinearidade entre as variáveis que compõem este sub-conjunto, pois dois dos Fatores de Inflação de Variância apresentam valores maiores que 10 (variáveis X_1 e X_2).

Modelo	Estatística de Colinearidade		
	Tolerância	VIF	
1	X01	,031	32,182
	X02	,031	32,644
	X03	,789	1,267
	X08	,344	2,906
	X15	,337	2,972
	RAIZ_X06	,895	1,117

Tabela 50: Estatísticas de Colinearidade

A Tabela 51 corrobora com os achados da tabela acima. Como pode ser observado na dimensão 7, que possui Índice de Condição maior que 15, as proporções de variância para as variáveis X_1 e X_2 são maiores que 0,90, indicando que estas variáveis estão altamente correlacionadas.

Modelo	Dimensão	Eigenvalue	Índice de Condição	Proporção de Variância						
				(Constante)	X01	X02	X03	X08	X15	RAIZ_X06
1	1	3,443	1,000	,00	,00	,00	,01	,01	,01	,01
	2	1,674	1,434	,00	,00	,00	,02	,03	,02	,01
	3	1,468	1,532	,00	,00	,00	,00	,07	,05	,00
	4	,206	4,089	,00	,00	,00	,18	,35	,43	,14
	5	,157	4,682	,00	,00	,00	,31	,52	,41	,17
	6	3,870E-02	9,432	,96	,00	,01	,49	,03	,01	,65
	7	1,360E-02	15,911	,03	,99	,98	,01	,00	,07	,01

Tabela 51: Eigenvalues, índice de condição e proporções de variância.

Em virtude destes resultados, a opção adotada foi a de não considerar este sub-conjunto a fim de evitar grandes erros na estimação dos coeficientes da regressão.

4.2.2.2 Sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$

Os resultados apresentados nas Tabelas 52 e 53 mostram que para este sub-conjunto de variáveis não há problemas de multicolinearidade, e portanto ele será usado no modelo de regressão logística.

Modelo		Estatística de Colinearidade	
		Tolerância	VIF
1	X03	,833	1,200
	X08	,345	2,897
	X15	,360	2,774
	RAIZ_X06	,925	1,082

Tabela 52: Estatísticas de Colinearidade.

Modelo	Dimensão	Eigenvalue	Índice de Condição	Proporção de Variância				
				(Constante)	X03	X08	X15	RAIZ_X06
1	1	3,012	1,000	,01	,01	,01	,01	,01
	2	1,582	1,380	,00	,01	,09	,07	,00
	3	,206	3,826	,00	,19	,34	,46	,15
	4	,158	4,361	,00	,33	,54	,43	,17
	5	4,213E-02	8,456	,99	,46	,02	,03	,67

Tabela 53: Eigenvalues, índice de condição e proporções de variância.

4.2.2.3 Sub-conjunto $[X_1; X_2; X_3; X_8; X_9; X_{11}; X_{12}; X_{15}; X_{16}]$

De acordo com os dados da Tabela 54 a seguir, há sérios problemas de multicolinearidade entre os dados que compõem este sub-conjunto. Por conta disto, ele também será descartado do modelo.

Modelo	Estatística de Colinearidade		
	Tolerância	VIF	
1	X01	,031	32,554
	X02	,030	33,044
	X03	,593	1,686
	X08	,324	3,085
	X09	,000	2868,056
	X11	,031	32,372
	X12	,000	3334,090
	X15	,284	3,527
	RAIZ_X06	,568	1,759

Tabela 54: Estatísticas de Colinearidade.

4.2.2.4 Sub-conjunto $[X_3; X_{14}; X_{r6}]$

Os resultados apresentados nas duas tabelas a seguir não apresentam nenhum indicativo de que haja problemas de multicolinearidade entre as variáveis que compõem o sub-conjunto. Portanto, ele será utilizado dentro do modelo de regressão logística.

Modelo	Estatística de Colinearidade		
	Tolerância	VIF	
1	X03	,980	1,021
	X14	,987	1,013
	RAIZ_X06	,992	1,008

Tabela 55: Estatísticas de Colinearidade.

Modelo	Dimensão	Eigenvalue	Índice de Condição	Proporção de Variância			
				(Constante)	X03	X14	RAIZ_X06
1	1	3,131	1,000	,01	,02	,03	,01
	2	,634	2,223	,00	,03	,90	,01
	3	,184	4,124	,01	,58	,03	,32
	4	5,091E-02	7,842	,98	,37	,04	,66

Tabela 56: Eigenvalues, índice de condição e proporções de variância.

4.2.3 Desenvolvimento do Modelo Logístico

A fim de desenvolver o modelo de Regressão Logística para classificação de empresas nos grupos de solventes e insolventes, será utilizado como critério de validação a separação da amostra original em amostra de classificação e amostra de

validação. Uma vez que a quantidade de dados disponíveis não é grande, esta divisão pode prejudicar o modelo desenvolvido para classificação. A fim de evitar que grandes problemas ocorram, e também por que não há um consenso geral sobre que proporção da amostra inicial deve ser destacada para validação, serão utilizadas 3 diferentes proporções da amostra inicial para validação. A primeira amostra de validação, a ser chamada de *Selec80* é composta de 80% da amostra inicial para desenvolvimento do modelo e de 20% para validação. A amostra *Selec60* contará com 40% da amostra inicial para validação e a *Selec50* conta com o mesmo número de casos tanto para a amostra de desenvolvimento do modelo como para a amostra de validação.

O método de seleção dos casos em cada um dos grupos se deu por meio de sorteio pseudo-aleatório, utilizando-se do procedimento de geração de números aleatórios do pacote estatístico SPSS, versão 9.0.0.

4.2.3.1 Modelo Logístico para o sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$

Para analisar o modelo logístico usando o sub-conjunto de variáveis $[X_3; X_8; X_{15}; X_{r6}]$, foram desenvolvidos os 3 modelos usando-se as amostras *Selec80*, *Selec60* e *Selec50*. O resumo dos resultados encontrados pode ser visto na tabela a seguir, onde podem ser vistos dois testes de adequação do modelo logístico (Qui-quadrado e o teste de Hosmer-Lemeshow) e os resultados preditos corretamente alcançados tanto pela amostra de classificação como para a amostra de validação do modelo. O resultado do teste de qui-quadrado de adequação do modelo logístico deve se mostrar significativo para que o modelo possa ser considerado um bom modelo de classificação. Já o teste

de adequação de Hosmer-Lemeshow mede a correspondência entre os valores reais e preditos da variável dependente, e em caso de boa adequação, o p-valor não deve se mostrar significativo, mostrando que há pequena diferença entre a classificação predita e a classificação observada.

	<i>Selec80</i>	<i>Selec60</i>	<i>Selec50</i>
Qui-quadrado (<i>Goodness-of-fit</i> do Modelo)	16,010	16,044	10,332
p-valor	0,0030	0,0030	0,0352
Hosmer-Lemeshow (p-valor)	0,6026	0,7825	0,1034
Amostra de Classificação (acerto)			
Insolventes	75%	75%	80%
Solventes	87,5%	100%	90%
Geral	81,25%	87,5%	85%
Amostra de Validação (acerto)			
Insolventes	75%	87,5%	80%
Solventes	75%	50%	70%
Geral	75%	68,75%	75%
Ambas as Amostras (acerto)	80%	80%	80%

Tabela 57: Resultados dos modelos logísticos usando o sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$.

É possível observar que o modelo gerado a partir do conjunto de dados *Selec50* apresenta resultados ligeiramente inferiores aos outros dois modelos quando se compara os resultados dos dois testes de adequação. No entanto, com relação ao índice de acerto de classificações, essa deficiência não se mostra presente, e é este o modelo que maior índice de acerto possui para empresas insolventes usando a amostra de classificação.

Os coeficientes das três equações de regressão usadas na etapa de classificação das empresas entre solventes e insolventes estão apresentados na tabela a seguir, juntamente com seus respectivos níveis de significância.

	X_{03}	X_{08}	X_{15}	X_{r06}	Constante
<i>Select80</i>	-4,3116**	1,0340	10,6971***	4,3308*	-2,1879
<i>Select60</i>	-7,2733***	1,6796	6,2464	7,0814**	-2,9456
<i>Select50</i>	-7,5470**	1,0841	5,1849	6,0054**	-1,4878

* Significante ao nível de 5%

** Significante ao nível de 10%

*** Significante ao nível de 15%

Tabela 58: Coeficientes dos modelos de regressão para o sub-conjunto $[X_3; X_8; X_{15}; X_{r6}]$.

4.2.3.2 Modelo Logístico para o sub-conjunto $[X_3; X_{14}; X_{r6}]$

Os resultados do modelo logístico desenvolvido para o sub-conjunto de variáveis $[X_3; X_{14}; X_{r6}]$ estão mostrados na tabela a seguir.

	<i>Sele80</i>	<i>Sele60</i>	<i>Sele50</i>
Qui-quadrado (<i>Goodness-of-fit</i> do Modelo)	11,235	19,876	22,282
p-valor	0,0105	0,0028	0,0240
Hosmer-Lemeshow (p-valor)	0,2275	0,4519	0,3426
Amostra de Classificação (acerto)			
Insolventes	68,75%	75%	70%
Solventes	93,75%	91,67%	100%
Geral	81,25%	83,33%	85%
Amostra de Validação (acerto)			
Insolventes	100%	87,5%	80%
Solventes	75%	50%	70%
Geral	87,5%	68,75%	75%
Ambas as Amostras (acerto)	82,5%	77,5%	80%

Tabela 59: Resultados dos modelos logísticos usando o sub-conjunto $[X_3; X_{14}; X_{r6}]$.

Também com este sub-conjunto de variáveis é possível perceber que os três modelos gerados apresentam resultados de predição extremamente semelhantes. O modelo gerado a partir da amostra *Select80* é ligeiramente superior aos demais com

relação à capacidade de previsão geral, e ainda foi o único modelo a acertar em 100% a predição de empresas insolventes durante a etapa de validação, mostrando ser um modelo muito útil para a previsão de empresas nesta situação devido a sua capacidade de generalização.

Os coeficientes das equações de regressão usadas na etapa de classificação estão apresentados na tabela a seguir.

	X_{03}	X_{14}	X_{r06}	Constante
<i>Select80</i>	-2,6923	-2,7746	4,6051*	-1,4755
<i>Select60</i>	-4,6555	-2,0057	7,2501*	-2,9294
<i>Select50</i>	-4,8087**	-1,3070	5,9819***	-1,8407

* Significante ao nível de 5%

** Significante ao nível de 10%

*** Significante ao nível de 15%

Tabela 60: Coeficientes dos modelos de regressão para o sub-conjunto [X_3 ; X_{14} ; X_{r6}].

4.3. Redes Neurais Artificiais

Uma vez que o tipo de RNA mais comumente encontrado na literatura para problemas desta natureza é o de retro-propagação, este foi escolhido também para o presente estudo. Todas as redes construídas contaram com uma camada de entrada, uma camada intermediária e uma camada de saída. Inicialmente foram adotados 4 neurônios na camada intermediária. Uma análise de sensibilidade feita a fim de encontrar o número de neurônios que fornece a menor média de MSE, possibilitou encontrar uma quantidade menor e ainda adequada para o conjunto de dados em questão. O número máximo de ciclos (*epochs*) foi limitado em 1.000, com uma regra

de interrupção que fez com que o processamento parasse quando o valor de MSE ficou menor que 0,01 para o conjunto de dados de validação. Todas as redes foram treinadas 10 vezes, onde em cada uma destas rodadas foram atribuídos aleatoriamente pesos iniciais para as conexões entre os neurônios.

Para a construção do modelo de RNA, fez-se uso do mesmo critério de divisão da amostra utilizado na construção do modelo logístico, ou seja, dividiram-se os dados em três diferentes proporções a fim de treinar e testar a rede. As proporções dos dados totais adotadas para treinar a rede foram de 80% (RNA-80), 60% (RNA-60) e 50% (RNA-50). Será usando também em cada um dos casos 10% do total dos casos como critério de parada do treinamento, garantindo que a rede em treinamento seja útil para generalizações. O percentual restante será usado para testar a rede.

Inicialmente, foram utilizadas para construir a rede todas as 16 variáveis mais as duas variáveis (X_{04} e X_{06}) transformadas por suas respectivas raízes quadradas. A opção pela quantidade total de variáveis disponíveis se deu para que não fossem descartadas inicialmente quaisquer delas que pudessem ser úteis na predição. No entanto, durante a análise de sensibilidade da rede a expectativa é de que esse número venha a ser reduzido sem representar perda para os resultados de predição do modelo.

4.3.1 Rede com Treinamento de 80% dos Casos

A fase inicial do processamento da rede é a fase de treinamento. Nesta etapa os dados são apresentados à rede que aprende e se adapta para conseguir prever

adequadamente todos os casos do problema. O maior indicador de que o processo de aprendizagem se deu de forma adequada é através da análise do MSE médio das 10 rodadas de treinamento. O gráfico abaixo apresenta o comportamento desta variável ao longo dos ciclos de aprendizagem.

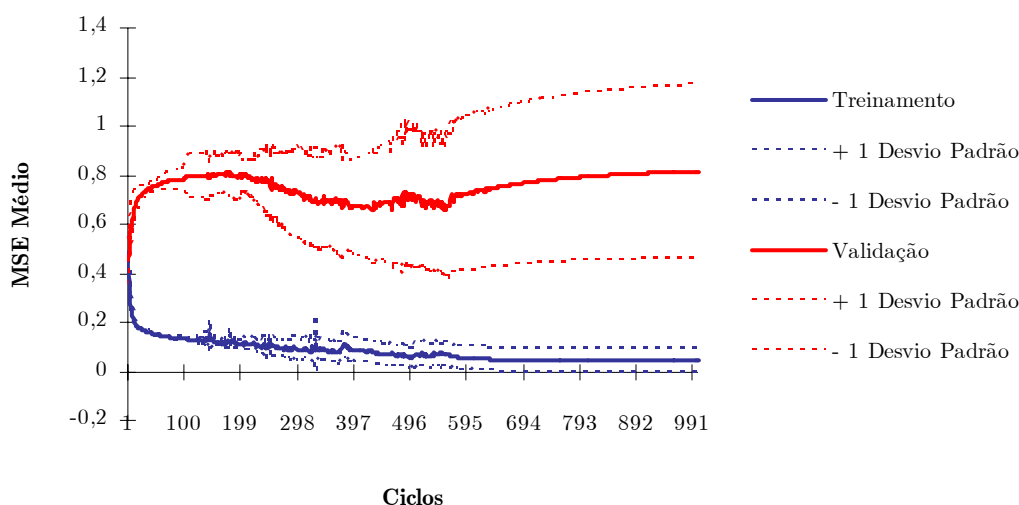


Gráfico 11: MSE médio das 10 rodadas para as amostras de treinamento e validação ao longo dos ciclos de aprendizagem da RNA-80.

Pode-se perceber que a média dos erros médios ao quadrado para a amostra de treinamento cai continuamente, convergindo para um valor bem próximo de zero. No entanto, para a amostra de validação esse valor não converge igualmente, como seria conveniente. Isto quer dizer que haverá problemas de classificação para a amostra de validação, uma vez que a rede não se mostrou tão generalizável.

A tabela abaixo apresenta os dados mais importantes encontrados no gráfico acima. É possível ver claramente que o valor mínimo obtido para a média dos MSEs com a amostra de validação foi pouco menor que 0,40, enquanto que para a amostra

de treinamento foi pouco menor que 0,035. É importante ressaltar que 0,40 é um valor muito alto para o erro, já que os dados de saída variam entre 0 e 1.

<i>Todas as Rodadas</i>	<i>Mínimo (Treinamento)</i>	<i>Desvio Padrão (Treinamento)</i>	<i>Mínimo (Validação)</i>	<i>Desvio Padrão (Validação)</i>
Média dos MSEs Mínimos	0,034500897	0,04127758	0,396873862	0,058300477
Média dos MSEs Finais	0,043397177	0,04731667	0,815621316	0,354527831

Tabela 61: Valores mínimos e desvio padrão para o MSE médio das 10 rodadas para as amostras de treinamento e validação da RNA-80.

Os dados da Tabela 62 apresentam os resultados da classificação das empresas em cada grupo para cada um dos conjuntos de dados (treinamento - 80% dos casos, validação - 10% dos casos e Teste - 10% restante). É possível perceber que a rede foi capaz de classificar corretamente 100% dos casos usados na amostra de aprendizagem. No entanto, tanto para a amostra de validação como para a amostra de teste, 1 caso foi classificado incorretamente, obtendo-se um percentual de 75% de classificação correta em ambos os casos.

Para se saber se o número de 4 neurônios na camada intermediária é uma quantidade adequada para a rede, foi realizada uma análise de sensibilidade a fim de encontrar a quantidade que resulta em uma menor média dos MSEs mínimos. A análise gráfica está mostrada no Gráfico 12.

		Grupo Predito			
		BKT	NBK	Total	
Treinamento	Contagem	BKT	16	0	16
		NBK	0	16	16
	%	BKT	100	0	100
		NBK	0	100	100
Validação	NBK	BKT	2	0	2
		NBK	1	1	2
	BKT	BKT	100	0	100
		NBK	50	50	100
Teste	NBK	BKT	1	1	2
		NBK	0	2	2
	BKT	BKT	50	50	100
		NBK	0	100	100

Tabela 62: Resultados de classificação para a RNA-80 contendo as 18 variáveis de predição.

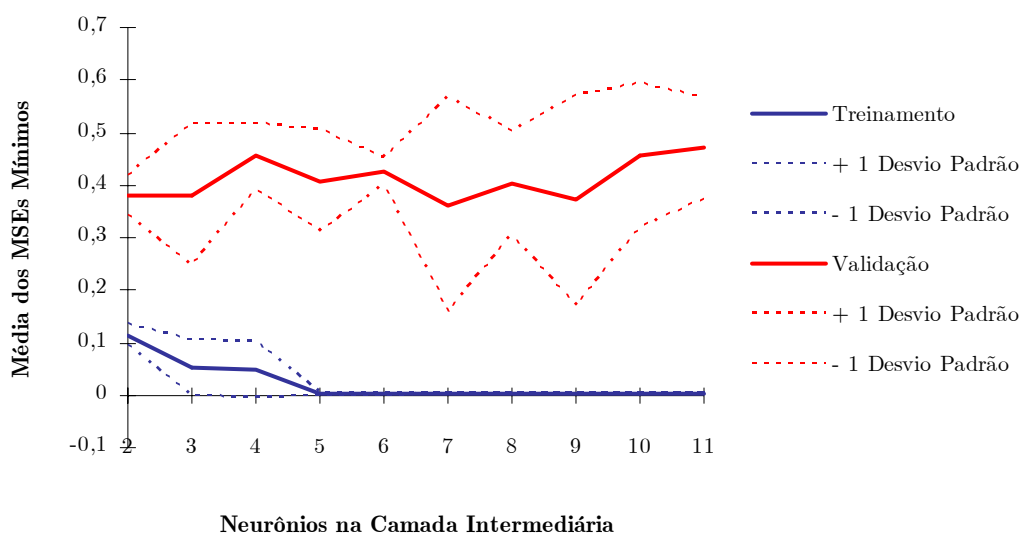


Gráfico 12: Média dos MSEs mínimos para diferentes quantidades de neurônios na camada intermediária.

É possível ver a partir do gráfico acima que a quantidade de neurônios ideal para a camada intermediária é cinco, pois a partir deste valor a média dos MSEs mínimos se mantém muito próxima a zero e não há mais reduções consideráveis. Uma

nova rede foi então criada com 5 neurônios na camada intermediária, mas nenhum ganho em classificação foi observado. Decidiu-se, portanto manter os resultados da rede com 4 neurônios.

Para descobrir quais das 18 variáveis utilizadas até agora são realmente importantes para o modelo, foi realizada uma análise de sensibilidade. Esta análise fornece uma medida da importância relativa dos dados de entrada da rede e mostra como o resultado de saída da rede se altera de acordo com a variação destas variáveis de predição. O processo de análise consiste em variar cada uma das variáveis entre sua média ± 1 desvio padrão, enquanto todas as outras variáveis são mantidas constantes em suas respectivas médias. O resultado da classificação é então obtido para 50 ciclos abaixo e acima de sua média. Este processo se repete para cada uma das variáveis disponíveis.

O gráfico a seguir apresenta um resumo da variação de cada dado de saída com relação à variação de cada variável de entrada. Desta análise é possível perceber que apenas as variáveis X_{03} , X_{14} , X_{r6} , X_{05} , e X_{16} deveriam fazer parte da rede. A variável X_{06} apresenta uma sensibilidade comparável à da variável X_{05} incluída na seleção, mas decidiu-se por não escolhê-la para fazer parte do grupo por ser altamente correlacionada com a variável X_{r6} .

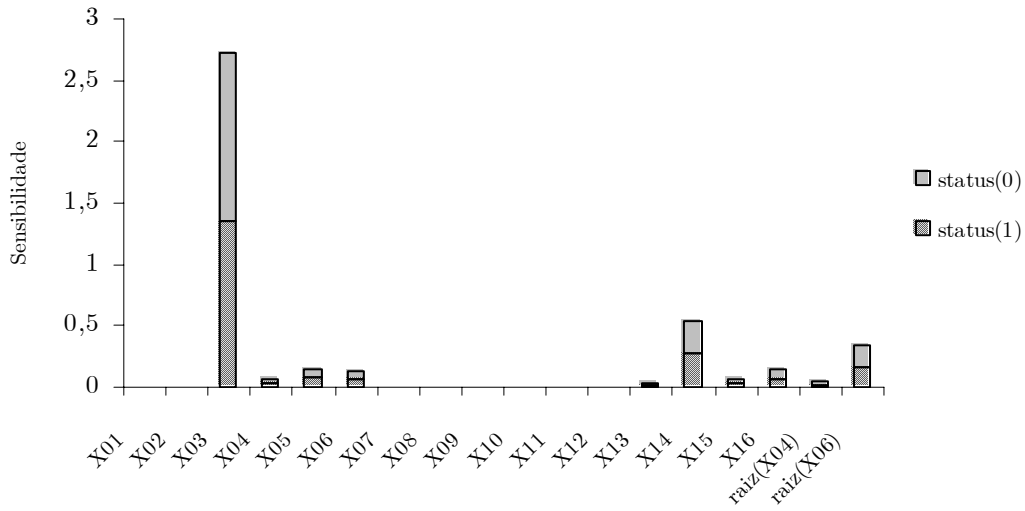


Gráfico 13: Análise de sensibilidade das variáveis do modelo RNA-80.

A fim de testar se uma rede com menos variáveis se comporta tão bem quanto a rede anterior formada pelas 18 variáveis, todo o processo de treinamento da nova rede foi realizado. Os dados apresentados na tabela abaixo mostram que a média dos MSEs mínimos e finais da amostra de treinamento melhoraram bastante em comparação com a rede original e que a média dos MSEs finais para a amostra de validação também apresentou melhora considerável.

<i>Todas as Rodadas</i>	<i>Mínimo (Treinamento)</i>	<i>Desvio Padrão (Treinamento)</i>	<i>Mínimo (Validação)</i>	<i>Desvio Padrão (Validação)</i>
Média dos MSEs Mínimos	0,006500388	0,00531327	0,419312418	0,022284966
Média dos MSEs Finais	0,006500388	0,00531327	0,667969108	0,187841073

Tabela 63: Valores mínimos e desvio padrão para o MSE médio das 10 rodadas para as amostras de treinamento e validação da rede com 5 neurônios e 5 variáveis de entrada.

Esta nova rede apresentou resultados de classificação muito semelhantes às redes anteriores. Apenas para a amostra de classificação esta rede piorou seu

resultado, classificando incorretamente 1 empresa a mais que a rede original. O resumo dos dados está mostrado na Tabela 64.

Uma vez obtidos todos estes resultados apresentados até agora, é possível seguramente afirmar que uma rede com apenas 5 variáveis e com 5 elementos de processamento na camada intermediária fornece bons resultados de classificação. Os resultados obtidos tanto para a amostra de validação como para a amostra de teste são altamente sensíveis às pequenas variações de classificação observadas dadas as pequenas quantidades de casos que fazem parte de suas respectivas amostras.

Uma análise mais robusta da validade dos procedimentos adotados para a construção das redes poderá ser feita nas seções seguintes, quando uma maior quantidade de dados foi usada na amostra de teste, garantindo uma melhor capacidade de generalização da rede.

		Grupo Predito		Total	
		BKT	NBK		
Treinamento	Contagem	BKT	16	0	16
		NBK	0	16	16
	%	BKT	100	0	100
		NBK	0	100	100
Validação	NBK	BKT	1	1	2
		NBK	1	1	2
	BKT	BKT	50	50	100
		NBK	50	50	100
Teste	NBK	BKT	2	0	2
		NBK	1	1	2
	BKT	BKT	100	0	100
		NBK	50	50	100

Tabela 64: Resultados de classificação para a RNA-80 contendo as 5 variáveis de predição.

4.3.2 Rede com Treinamento de 60% dos Casos

Nesta segunda etapa do processo de desenvolvimento da RNA, foram utilizados para compor a amostra do processo de aprendizagem 60% dos casos disponíveis, ou seja, 24 empresas. Das 16 empresas restantes 4 foram usadas como critério de parada no processo de treinamento (amostra de validação da rede) e as 12 empresas restantes foram usadas para testar a capacidade de generalização da rede criada.

Em primeiro lugar, foi criada a rede usando-se os mesmos parâmetros utilizados na rede da seção anterior. Em seguida esta rede foi treinada em 10 rodadas consecutivas, onde em cada uma delas os pesos iniciais de treinamento foram atribuídos de maneira aleatória.

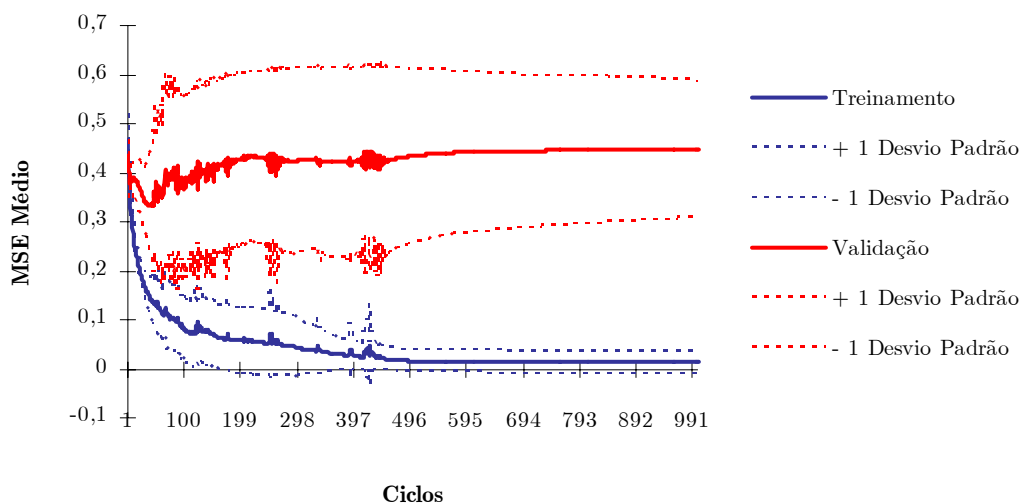


Gráfico 14: MSE médio das 10 rodadas para as amostras de treinamento e validação ao longo dos ciclos de aprendizagem da RNA-60.

Também para esta rede, a média dos erros médios ao quadrado para a amostra de treinamento caiu continuamente, aproximando-se bastante de zero. O mesmo não

ocorreu para a amostra de validação, que caiu um pouco no início dos ciclos mas depois se estabilizou em um valor em torno de 0,44. Isto mostra que para esse conjunto de 4 empresas haverá problemas de classificação.

A tabela abaixo apresenta o resumo dos dados mais importantes apresentados no gráfico acima. É possível notar que o valor mínimo obtido para a média dos MSEs com a amostra de validação foi cerca de 0,20, cerca de metade do valor encontrado na RNA-80 com todas as variáveis. Para a amostra de treinamento esse mesmo índice foi menor que 0,01, um valor também bastante inferior ao encontrado na RNA-80.

<i>Todas as Rodadas</i>	<i>Mínimo (Treinamento)</i>	<i>Desvio Padrão (Treinamento)</i>	<i>Mínimo (Validação)</i>	<i>Desvio Padrão (Validação)</i>
Média dos MSEs Mínimos	0,009908918	0,015353146	0,200540066	0,134046853
Média dos MSEs Finais	0,013775587	0,022514364	0,447839081	0,139679715

Tabela 65: Valores mínimos e desvio padrão para o MSE médio das 10 rodadas para as amostras de treinamento e validação da RNA-60.

A tabela a seguir apresenta os resultados de classificação das empresas. A rede foi capaz de classificar corretamente 87,5% de todas as empresas da amostra de treinamento. Para a amostra de validação, 100% das empresas foram classificadas corretamente, mesmo com a não convergência adequada do MSE médio. No entanto, para a amostra de teste, que mede a real capacidade de generalização da rede desenvolvida os dados não foram bons: apesar de 100% das empresas insolventes terem sido classificadas corretamente com tais, 83,33% das empresas efetivamente solventes foram classificadas como insolventes. Para o caso de detecção apenas de empresas em situação de insolvência, o modelo poderia ser considerado adequado,

mas traria embutido em si uma grande tendência a comprometer a classificação das empresas solventes. No caso de seu uso para análise de crédito, uma grande quantidade de bons negócios poderia ser perdida.

Como esta rede ainda conta com todas as 18 variáveis disponíveis, as relações entre as próprias variáveis e a existência de multicolinearidade podem ter influenciado sobremaneira os resultados de classificação da rede. A análise de sensibilidade para a seleção das variáveis e também para a identificação da quantidade adequada de neurônios na camada intermediária devem melhorar os resultados de classificação.

		Grupo Predito		Total	
		BKT	NBK		
Treinamento	Contagem	BKT	11	1	12
		NBK	2	10	12
	%	BKT	91,67	8,33	100
		NBK	16,67	83,33	100
Validação	NBK	BKT	2	0	2
		NBK	0	2	2
	BKT	BKT	100	0	100
		NBK	0	100	100
Teste	NBK	BKT	6	0	6
		NBK	5	1	6
	BKT	BKT	100,00	0,00	100
		NBK	83,33	16,67	100

Tabela 66: Resultados de classificação para a RNA-60 contendo as 18 variáveis de predição.

O gráfico abaixo mostra que a quantidade de neurônios adequada para a camada intermediária é 4.

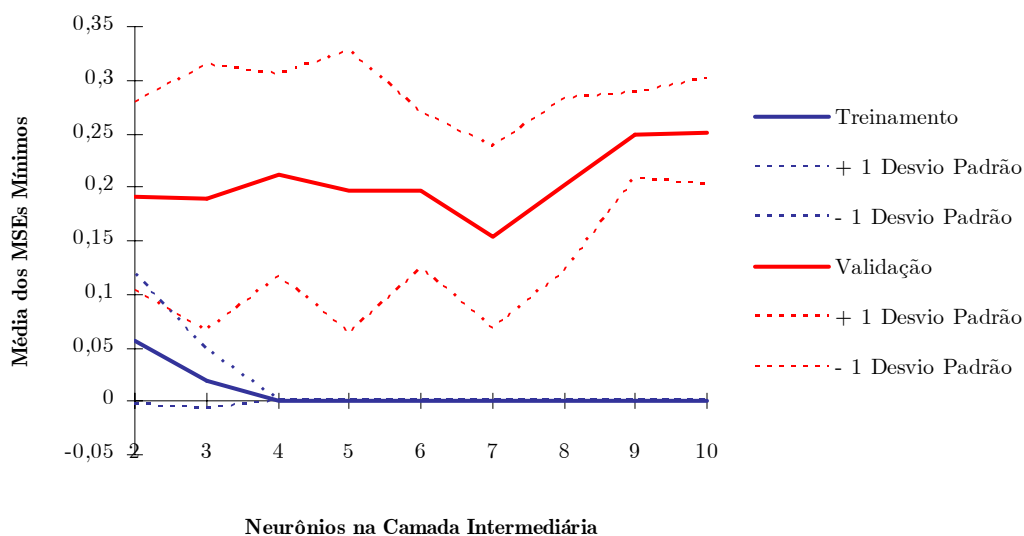


Gráfico 15: Média dos MSEs mínimos para diferentes quantidades de neurônios na camada intermediária.

A análise de sensibilidade indicou que as variáveis X_{11} , X_{r6} , X_{07} , X_{12} , X_{05} , X_{09} e X_{r4} devem compor a rede. A variável X_{06} apresentou uma sensibilidade inclusive maior que a da variável X_{r4} , mas não foi incluída na seleção por ser altamente correlacionada com a variável X_{r6} , que apresentou uma maior importância para a rede.

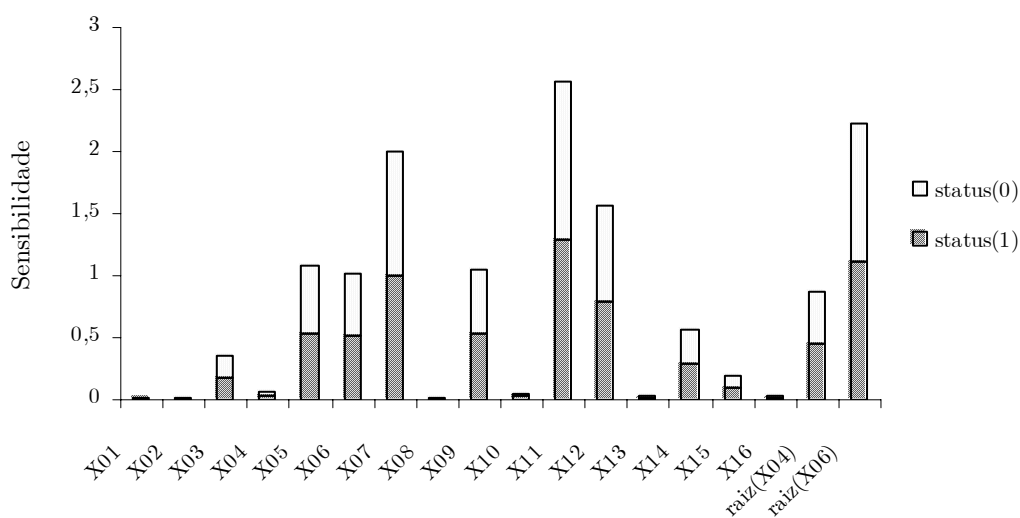


Gráfico 16: Análise de sensibilidade das variáveis do modelo RNA-60.

Para testar a nova rede com menos variáveis, mas ainda com 4 neurônios na camada intermediária, todo o processo de treinamento da nova rede foi realizado. Esta nova rede apresentou resultados de classificação muito melhores que a rede com 18 variáveis durante a fase de teste, o que mostra que ela possui uma capacidade de generalização muito superior à anterior. Na fase de treinamento a rede também melhorou a sua capacidade de classificação, atingindo um índice total de acerto de 95,83% dos casos. Para os dados de validação, a rede continuou acertando em 100% dos casos.

		Grupo Predito		Total	
		BKT	NBK		
Treinamento	Contagem	BKT	11	1	12
		NBK	0	12	12
	%	BKT	91,67	8,33	100
		NBK	0,00	100,00	100
Validação	NBK	BKT	2	0	2
		NBK	0	2	2
	BKT	BKT	100	0	100
		NBK	0	100	100
Teste	NBK	BKT	5	1	6
		NBK	1	5	6
	BKT	BKT	83,33	16,67	100
		NBK	16,67	83,33	100

Tabela 67: Resultados de classificação para a RNA-60 contendo 7 variáveis de predição.

A partir destes achados, é possível afirmar que a rede com apenas 7 variáveis e com 4 neurônios na camada intermediária fornece resultados de classificação confiáveis e generalizáveis. Estes resultados são inclusive muito superiores a todos os

outros resultados encontrados até o momento. A seção seguinte apresentará os resultados obtidos com a RNA-50, a última rede a ser construída neste experimento.

4.3.3 Rede com Treinamento de 50% dos Casos

Nesta etapa em que foram usadas 50% das empresas para compor o grupo de treinamento, os resultados para o MSE foram ainda melhores do que os obtidos com a amostra de treinamento da RNA-60. Nesta nova rede, a média dos MSEs mínimos para as 10 rodadas de treinamento foi menor que 0,0003, resultado obtido no último ciclo de aprendizagem. A média dos MSEs mínimos para a amostra de validação foi próxima, mas ligeiramente superior ao resultado obtido com a RNA-60, alcançando 0,22. O gráfico e a tabela abaixo apresentam estes dados.

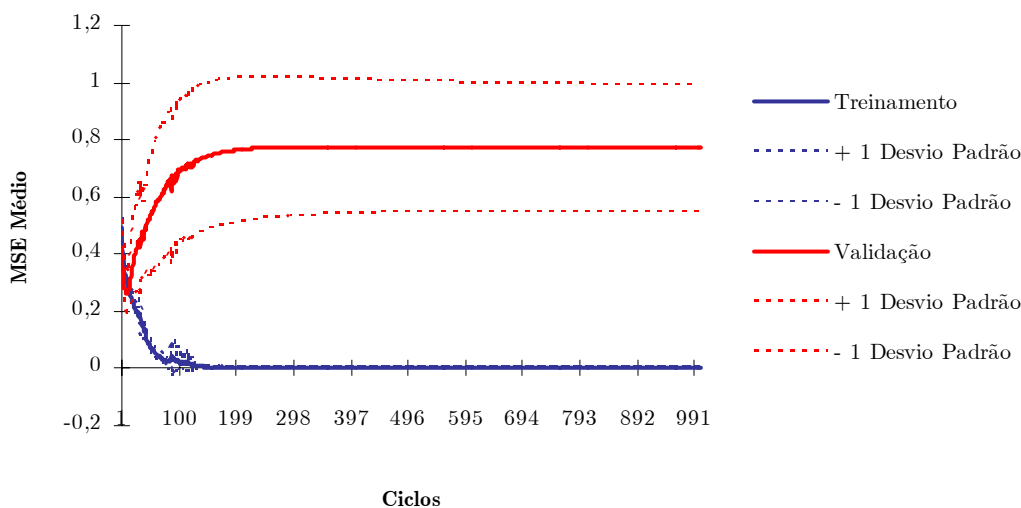


Gráfico 17: MSE médio das 10 rodadas para as amostras de treinamento e validação ao longo dos ciclos de aprendizagem da RNA-50.

<i>Todas as Rodadas</i>	<i>Mínimo (Treinamento)</i>	<i>Desvio Padrão (Treinamento)</i>	<i>Mínimo (Validação)</i>	<i>Desvio Padrão (Validação)</i>
Média dos MSEs Mínimos	0,000267413	0,000178933	0,224409923	0,078579754
Média dos MSEs Finais	0,000267413	0,000178933	0,769518852	0,218311653

Tabela 68: Valores mínimos e desvio padrão para o MSE médio das 10 rodadas para as amostras de treinamento e validação da RNA-50.

A tabela seguinte mostra que a RNA-50 com as 18 variáveis não apresentou resultados de classificação satisfatórios. Apenas 65% dos casos foram classificados corretamente na própria amostra de aprendizagem. Na amostra de validação, todos os 4 casos foram classificados corretamente, e na amostra de teste apenas 56,25% dos casos classificados foram acertados. Mais uma vez a rede classificou durante a fase de teste as empresas insolventes melhor do que as empresas solventes.

		Grupo Predito			
			BKT	NBK	Total
Treinamento	Contagem	BKT	8	2	10
		NBK	5	5	10
	%	BKT	80,00	20,00	100
		NBK	50,00	50,00	100
Validação	NBK	BKT	2	0	2
		NBK	0	2	2
	BKT	BKT	100	0	100
		NBK	0	100	100
Teste	NBK	BKT	8	0	8
		NBK	7	1	8
	BKT	BKT	100,00	0,00	100
		NBK	87,50	12,50	100

Tabela 69: Resultados de classificação para a RNA-50 contendo as 18 variáveis de predição.

Para se saber se a quantidade de neurônios na camada intermediária deveria se manter em 4, a análise de sensibilidade para este parâmetro foi realizada. O gráfico a seguir apresenta os resultados obtidos.

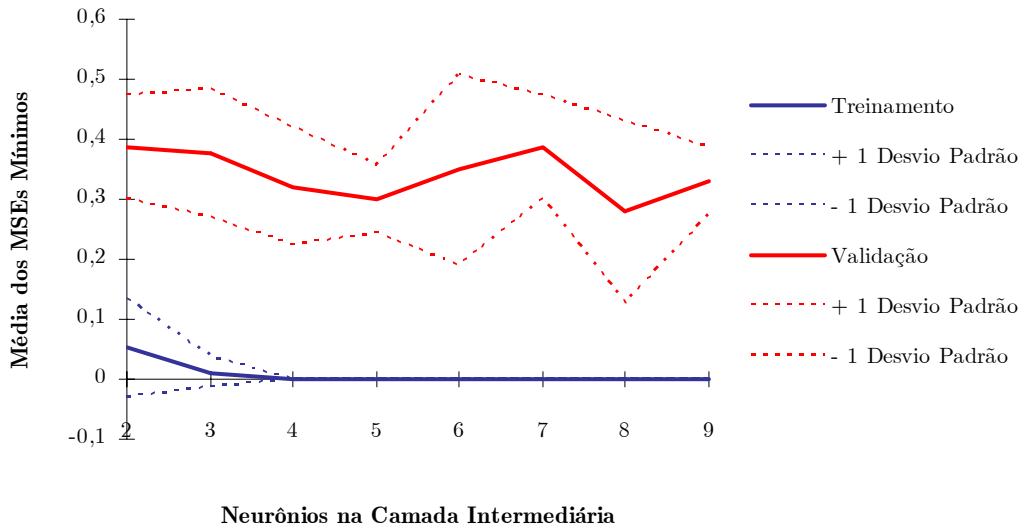


Gráfico 18: Média dos MSEs mínimos para diferentes quantidades de neurônios na camada intermediária.

É possível perceber que seria possível testar toda a rede com apenas 3 neurônios, uma vez que a média dos MSEs mínimos já se aproxima bastante de zero com esta quantidade. No processo de melhoria da qualidade de predição da rede, a análise de sensibilidade mostrou que apenas as variáveis X_{03} , X_{07} , X_{09} , X_{14} , X_{16} , X_{13} e X_{12} não necessitam ser dispensadas do modelo. O gráfico seguinte apresenta estes resultados.

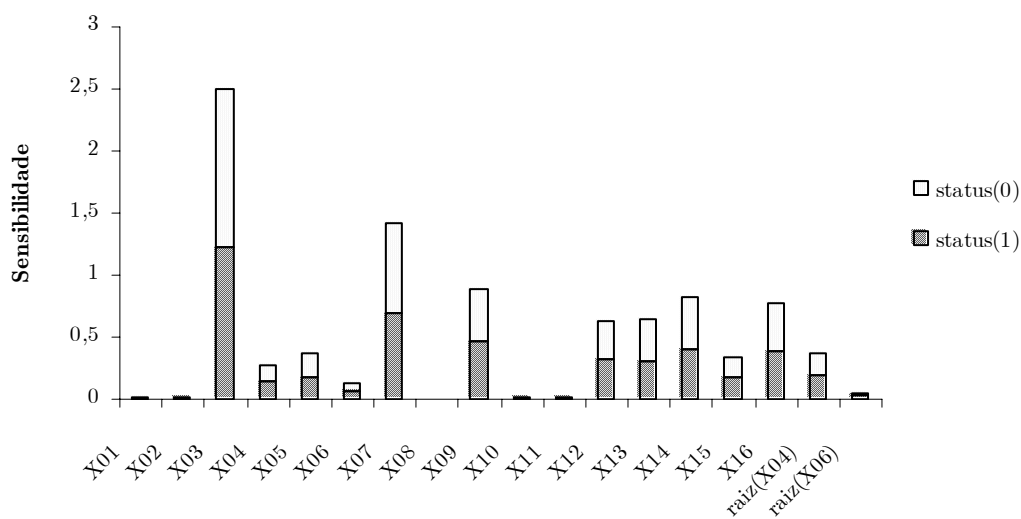


Gráfico 19: Análise de sensibilidade das variáveis do modelo RNA-50.

Para testar a nova configuração de rede com menos variáveis, foram construídas 2 novas redes: a primeira com 3 elementos na camada intermediária e a segunda ainda com 4 elementos. A nova rede com 3 elementos apresentou classificações iguais às da rede com 4 elementos para as amostras de validação e de teste. No entanto, para a amostra de treinamento, a rede com 3 neurônios teve um desempenho bastante inferior (75% *versus* 100%). Os resultados de classificação para a rede com 4 neurônios são mostrados na tabela a seguir.

		Grupo Predito			
		BKT	NBK	Total	
Treinamento	Contagem	BKT	10	0	10
		NBK	0	10	10
	%	BKT	100	0	100
		NBK	0	100	100
Validação	NBK	BKT	2	0	2
		NBK	0	2	2
	BKT	BKT	100	0	100
		NBK	0	100	100
Teste	NBK	BKT	5	3	8
		NBK	1	7	8
	BKT	BKT	62,5	37,5	100
		NBK	12,5	87,5	100

Tabela 70: Resultados de classificação para a RNA-50 contendo 7 variáveis de predição.

5. Conclusões e Recomendações

O objetivo desta pesquisa foi o de desenvolver modelos de predição de insolvência utilizando dados brasileiros através da Análise Discriminante Linear, Regressão Logística e Redes Neurais Artificiais. O estudo foi conduzido com rigorosa atenção aos pré-requisitos das técnicas, a fim de não criar modelos estatisticamente inadequados. Uma especial atenção foi dedicada ao desenvolvimento do modelo discriminante, por ser esta a técnica que mais impõe restrições à pesquisa.

Diferentemente da grande maioria dos estudos que tratam deste tema, o critério de seleção de variáveis não se deu por meio da técnica *stepwise*, mas sim pela técnica de melhores sub-conjuntos, uma metodologia mais moderna e que possibilita analisar uma maior quantidade de modelos e de forma mais rápida que a *stepwise*.

As variáveis usadas como candidatas a serem incluídas nos modelos foram divididas em 4 grupos. O primeiro grupo contou com variáveis que descreviam a estrutura de capital das empresas. O segundo grupo era formado por indicadores de liquidez, o terceiro grupo possuía variáveis de rentabilidade, e por fim, o quarto grupo, era formado por variáveis de atividade e rotação.

Das 3 variáveis de estrutura de capital, apenas a variável X_{03} , composição do endividamento, se mostrou útil para os modelos desenvolvidos. Por sinal, esta variável foi a mais presente durante o estudo, fazendo parte de 5 dos 7 modelos finais gerados. O mais importante ainda é comprovar a hipótese inicial de que esta variável deveria possuir um sinal negativo, indicando que quanto maior fosse o endividamento

de curto prazo maior seria a probabilidade da empresa ser classificada como insolvente. Esta hipótese foi confirmada em todos os modelos gerados.

Havia 3 variáveis dentro do grupo de indicadores de liquidez das empresas. A variável X_{05} , liquidez corrente, fez parte de 4 modelos. Ela foi incluída nos dois modelos discriminantes e em duas das redes neurais criadas. Vale lembrar que nenhuma variável fez parte dos três modelos de redes neurais criados. Da análise dos dados foi possível perceber que as empresas solventes eram mais líquidas que as insolventes, confirmando a hipótese inicial.

Outra variável importante nos modelos foi a raiz da variável X_{06} , liquidez seca. Esta variável fez parte dos dois modelos logísticos e de duas redes neurais, sempre coerente com a hipótese inicial de que quanto mais líquida a empresa maior a probabilidade de pertencer ao grupo de solventes.

Dentre as 6 variáveis de rentabilidade não houve uma que se destacasse individualmente fazendo parte de vários modelos. De todas elas apenas a variável X_{10} , retorno sobre o patrimônio líquido, não fez parte de modelo algum. Das outras cinco, três foram importantes em duas das três redes neurais desenvolvidas. Foram elas as variáveis X_{07} , giro do ativo, X_{09} , retorno sobre o ativo, e X_{12} , lucros retidos sobre o ativo.

O único grupo em que todas as suas variáveis fizeram parte de pelo menos um dos modelos foi o dos índices de atividade. Das 4 variáveis do grupo, a variável X_{14} , recebimento de vendas, foi a que esteve presente em mais modelos: três. A outra variável deste grupo que merece destaque é a X_{16} , estoque sobre venda, que foi utilizada em duas redes neurais.

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 foi a técnica mais pobre em termos de resultados alcançados. A validação de seus dados só alcançou 70% do total dos casos. Este valor é considerado ainda pior por se saber que para esta técnica não houve validação com dados fora da amostra, mas com a aplicação da técnica de validação do tipo *leave-one-out*.

A Regressão Logística apresentou resultados superiores. Para esta técnica, parte da amostra foi utilizada para a construção do modelo e o restante foi utilizado para validação. Para os dados de construção do modelo o índice de classificação correta variou entre 81,25% e 87,5%. A validação apresentou resultados entre 68,75% e 87,5%. Ao contrario da ADL, aqui a validação se deu com um grupo de empresas fora da amostra de construção do modelo, o que valoriza os resultados alcançados.

As três redes neurais construídas alcançaram resultados ainda melhores. Nesta técnica, contando-se as empresas que fizeram parte da amostra de treinamento (construção da rede) e de validação, entre 90% e 95% das empresas foram classificadas corretamente. Em apenas uma das redes a amostra de treinamento não alcançou 100% de acerto, alcançando 95,8%, resultados que mostram como as redes facilmente aprendem os dados e os generaliza para outros casos.

Em resumo, este trabalho mostrou que há métodos eficientes disponíveis para predição de situação de dificuldades financeiras de empresas e que estas técnicas têm evoluído sobremaneira ao longo dos anos, seja com relação ao método em si, à capacidade computacional disponível ou até mesmo quanto ao critério de seleção de variáveis.

Apesar de tudo aquilo que se comenta a respeito da qualidade das informações disponíveis nos demonstrativos financeiros das empresas, muitas análises eficazes podem ser realizadas desde que a técnica mais apropriada seja utilizada e que seus requisitos sejam rigorosamente atendidos.

A não regularidade de inclusão de indicadores nos modelos gerados mostrou que não há uma base teórica sólida com relação ao tema e que para se construir um bom modelo é necessário experimentar uma grande quantidade deles a fim de encontrar aqueles que fornecem os melhores indicativos da situação que se deseja prever. Este trabalho não buscou esgotar o tema, mas oferecer uma contribuição com dados atualizados e técnicas modernas. O maior interesse desta pesquisa é contribuir para esta linha de pesquisa e que desperte em outros pesquisadores o interesse por corroborar ou não os resultados obtidos.

Ficam como recomendações a pesquisas futuras a utilização de novas técnicas disponíveis para classificação, bem como também de novos critérios para seleção de variáveis. Viu-se durante a revisão da literatura disponível sobre o tema que o *stepwise* ainda é o critério mais largamente aplicado para a seleção de variáveis nos modelos usados para classificação. Isto ocorre apesar de já haver disponível tanto literatura quanto rotinas implementadas em aplicativos estatísticos para a adoção de critérios alternativos ao *stepwise*. Novos estudos podem também testar a utilidade para a predição de outras variáveis não utilizadas nesta pesquisa, uma vez que uma infinidade de indicadores pode ser criada a partir dos dados publicados pelas empresas. O que se buscou neste trabalho foi selecionar as variáveis mais comuns na literatura e que possuem entendimento mais claro. Por fim, é importante dizer que

novas pesquisas nesta linha ajudarão no desenvolvimento e divulgação do tema abordado e servirão para consolidar sua relevância acadêmica.

6. Referências Bibliográficas

ALMEIDA, F.C. *Desvendando o Uso de Redes Neurais em Problemas de Administração de Empresas*. Revista de Administração de Empresas, São Paulo, v.35, n.1, pp.46-55, janeiro/fevereiro de 1995.

_____ & DUMONTIER, P. *O Uso de Redes Neurais em Avaliação de Risco de Inadimplência*. Revista de Administração, v.31, n.1, p.52-63, Universidade de São Paulo, janeiro/março de 1996.

ALTMAN, E.I. *Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy*. Journal of Finance, v. 29, pp.589-609, 1968.

_____, HALDEMAN, R.G. & NARAYANAN, P. *ZetaTM Analysis: a new model to identify bankruptcy risk of corporations*. Journal of Banking and Finance, pp.29-54, 1977.

_____, BAYDIA, T.K.N. & DIAS, L.M.R. *Previsão de Problemas Financeiros em Empresas*. Revista de Administração de Empresas, 19(1), p.17-28. Fundação Getúlio Vargas, jan./mar., 1979.

_____. *Corporate Financial Distress: a complete guide to predicting, avoiding, and dealing with bankruptcy*. New York: John Wiley and Sons, 1983.

_____. *Corporate Financial Distress and Bankruptcy: a complete guide to predicting & avoiding distress and Profiting from Bankruptcy*. 2ª edição. New York: John Wiley and Sons, 1993.

_____, MARCO, G. & VARETTO, F. *Corporate Distress Diagnosis: comparisons using linear discriminant analysis and neural networks*. Journal of Banking and Finance, pp.505-29, 1994.

BACK, B., LAITINEN, T., SERE, K. & VAN WEZEL, M. *Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis and Genetic Algorithms*. Turku Centre for Computer Science, Technical Report número 40, setembro de 1996.

- BARROS, L.A.B.C. *Estrutura de Capital e Valor da Empresa no Contexto de Assimetria de Informações e Relações de Agência: um estudo empírico*. Dissertação de Mestrado, Faculdade de Economia Administração e Contabilidade. São Paulo: Universidade de São Paulo, 2001.
- BEAVER, W.H. *Financial Ratios as Predictors of Failure*. Journal of Accounting Research, janeiro de 1967.
- _____. *Market Prices, Financial Ratios, and the Prediction of Failure*. Journal of Accounting Research, pp.179-192, outono de 1968.
- BREALEY, R.A. & MYERS, S.C. *Princípios de Finanças Empresariais*. Portugal: Editora McGraw Hill de Portugal, 1992.
- CAOUILLE, J.B., ALTMAN, E.I. & NARAYANAN, P. *Gestão do Risco de Crédito: o próximo grande desafio financeiro*. Rio de Janeiro: Qualitymark, 1999.
- COATS, P.K. & FANT, L.F. *Recognizing Financial Distress Patterns Using a Neural Network Tool*. In TRIPPI, R.R. & TURBAN, E. (Eds.), *Neural Networks in Finance and Investing: using artificial intelligence to improve real-world performance*, Edição Revista, Estados Unidos: Irwin, pp.163-92, 1996.
- COCHRAN, W.G. *On the Performance of the Linear Discriminant Function*. Technometrics, vol.6, pp. 179-190, maio de 1964.
- DUN & BRADSTREET. *The Failure Record*. Nova Iorque: Dun & Bradstreet, 1997.
- DURAND, D.D. *Risk Elements in Consumer Installment Financing*. Studies in Consumer Installment Financing, National Bureau of Economic Research, pp. 105-142, 1941.
- FITZPATRICK, P. *A Comparison of the Ratios of Successful Industrial Enterprises with those of Failed Companies*. The Accountants Publishing Company, 1932.
- GILSON, S.C. *Managing Default: some evidence on how firms choose between workouts and* Capítulo 11. Journal of Applied Corporate Finance, verão de 1991.

- GILSON, S.C., JOHN, K. & LANG, L. *Troubled Debt Restructurings: An Empirical Study of Private Reorganization of Firms in Default*. Journal of Financial, pp. 315-354, outubro de 1990.
- GUJARATI, D.N. *Econometria Basica*. 3a. Edição. são Paulo: Makron Books, 2000.
- HÁFEZ, A. *Projeto que Dificulta Falência Causa Polemica*. In Gazeta Mercantil, São Paulo, 28 de outubro de 1996.
- HAIR, J.F., ANDERSON, R.E., TATHAM, R.L. & BLACK, W.C. *Multivariate Data Analysis with readings*. 5ª edição. New Jersey: Prentice Hall, 1998.
- HAWLEY, D.D., JOHNSON, J.D. & RAINA, D. *Artificial Neural Systems: a new tool for financial decision-making*. In TRIPPI, R.R. & TURBAN, E. (Eds.), *Neural Networks in Finance and Investing: using artificial intelligence to improve real-world performance*, Edição Revista, Estados Unidos: Irwin, pp.25-43, 1996.
- HOSMER, D.W. & LEMESHOW, S. *Applied Logistic Regression*. 2a. Edição, Estados Unidos: 2000.
- KANITZ, S.C. *Indicadores Contábeis e Financeiros de Previsão de Insolvência: a experiência na pequena e média empresa brasileira*. Tese (Livre Docência), Faculdade de Economia Administração e Contabilidade. São Paulo: Universidade de São Paulo, 1976.
- KAYO, E.K. & FAMÁ, R. *Dificuldades Financeiras, Custos de Agência e o Instituto Jurídico da Concordata*. Caderno de Pesquisas em Administração, v.1, n.3, São Paulo, 1996.
- KLIMASAUSKAS, C.C. *Applying Neural Networks*. In TRIPPI, R.R. & TURBAN, E. (Eds.), *Neural Networks in Finance and Investing: using artificial intelligence to improve real-world performance*, Edição Revista, Estados Unidos: Irwin, pp.45-69, 1996.
- LACHENBRUCH, P.A. *An Almost Unbiased Method of Obtaining Confidence Intervals for the Probability of Misclassification in Discriminant Analysis*. Biometrics, 23, pp.639-645, dezembro de 1967.
- _____, *Discriminant Analysis*. Hafner Press, 1975.

- LEI DAS FALÊNCIAS E CONCORDATAS. Decreto-Lei n. 7661 de 21 de junho de 1945.
- MATARAZZO, D.C. *Análise Financeira de Balanços: abordagem básica e gerencial*. 5ª. edição. São Paulo: Atlas, 1998.
- MATIAS, A.B. & SIQUEIRA, J.O. *Risco Bancário: modelo de previsão de insolvência de bancos no Brasil*. Revista de Administração da Universidade de São Paulo, vol.31, n.2, pp. 19-28, abril/junho de 1996.
- MEDSKER, L., TURBAN, E. & TRIPPI, R.R. *Neural Network Fundamentals for Financial Analysts* In TRIPPI, R.R. & TURBAN, E. (Eds.), *Neural Networks in Finance and Investing*, Estados Unidos: Probus Publishing Company, pp.3-25, 1993.
- MYERS, H. & FORGY, E.W. *Development of Numerical Credit Evaluation Systems*. Journal of American Statistical Association, vol.50, pp. 797-806, setembro de 1963.
- NETER, J., KUTNER, M.H., NACHTSHEIM, C.J. & WASSERMAN, W. *Applied Linear Statistical Models*. 4ª Edição, McGraw Hill, 1996.
- ODOM, M.D. & SHARDA, R. *A Neural Network Model for Bankruptcy Prediction*. In TRIPPI, R.R. & TURBAN, E. (Eds.), *Neural Networks in Finance and Investing*, Estados Unidos: Probus Publishing Company, pp.177-85, 1993
- OHLSON, J.A. *Financial Ratios and the Probabilistic Prediction of Bankruptcy*. Journal of Accounting Research, v.18, n.1, pp.109-131, primavera de 1980.
- PODDIG, T. *Bankruptcy Prediction: a comparison with discriminant analysis*. In Refenes, Apostolos-Paul (Ed.), *Neural Networks in the Capital Markets*, John Wiley and Sons, pp. 311-23, 1995.
- ROBINSON, P. *The idea of Performance Measurement: a working paper*. University of Alberta Working Paper, 19 de janeiro de 2002. Disponível on-line: <http://aux.zicklin.baruch.cuny.edu/critical/html2/8054robinson.html> em 15 de outubro de 2002.
- ROSS, S.A., WESTERFIELD, R.W. & JAFFE, J.F. *Administração Financeira: corporate finance*. 2ª Edição. São Paulo: Atlas, 2002.

- SANVICENTE, A.Z. & MINARDI, A.M.A.F. *Identificação de Indicadores Contábeis Significativos para a Previsão de Concordata de Empresas*. Finance Lab Working Papers, Ibmec Business School São Paulo. Outubro de 1998.
- SHLEIFER, A. & VISHNY, R.W. *A survey of Corporate Governance*. The Journal of Finance, v. LII, n. 2, pp. 737-83, junho de 1997.
- SILVA, J.P. *Análise Financeira das Empresas*. 5a. edição. Sao Paulo: Atlas, 2001.
- TASHJIAN, E., LEASE, R.C. & MCCONNELL, J.J. *Prepacks: an empirical analysis of prepackaged bankruptcies*. Journal of Financial Economics n.40, pp. 135-162, 1996.
- TRIPPI, R.R. & LEE, J.K. *Artificial Intelligence in Finance & Investing: state-of-the-art technologies for securities selection and portfolio management*. Chicago: Irwin, 1996.
- WALTER, J.E. *A Discriminant Function for Earnings Price Ratios of Large Industrial Corporations*. Review of Economics and Statistics, vol.41, pp. 44-52, fevereiro de 1959.
- WRUCK, K.H. *Financial Distress, reorganization, and organizational efficiency*. Journal of Financial and Economics 27, pp. 419-444, 1990.
- ZINGALES, L. *Corporate Governance*. The New Palgrave Dictionary of Economics and the Law. 1997.