

**UNIVERSIDADE FEDERAL DO PARANÁ**

**WYRLLEN EVERSON DE SOUZA**

**SISTEMAS MULTIAGENTES INTELIGENTES PARA RECONHECIMENTO E  
CLASSIFICAÇÃO DE PADRÕES APLICADOS EM PROBLEMAS DE SISTEMAS  
ELÉTRICOS DE POTÊNCIA**

**CURITIBA**

**2012**

**WYRLLEN EVERSON DE SOUZA**

**SISTEMAS MULTIAGENTES INTELIGENTES PARA RECONHECIMENTO E  
CLASSIFICAÇÃO DE PADRÕES APLICADOS EM PROBLEMAS DE SISTEMAS  
ELÉTRICOS DE POTÊNCIA**

Tese apresentada ao curso de Pós-Graduação em Métodos Numéricos em Engenharia, Área de Concentração em Programação Matemática, Departamento de Engenharia Civil, Departamento de Matemática dos setores de Tecnologia e de Ciências Exatas da Universidade Federal do Paraná, como requisito parcial à obtenção do título de Doutor em Métodos Numéricos em Engenharia.

Orientador: Prof. Dr. Anselmo Chaves Neto  
Co-Orientador: Prof. Dr. Alexandre Rasi Aoki

**CURITIBA**

**2012**

## TERMO DE APROVAÇÃO

WYRLLEN EVERSON DE SOUZA

### SISTEMAS MULTIAGENTES INTELIGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES APLICADOS EM PROBLEMAS DE SISTEMAS ELÉTRICOS DE POTÊNCIA

Tese aprovada como requisito parcial para obtenção do grau de Doutor no curso de Pós-Graduação em Métodos Numéricos em Engenharia, setores de Tecnologia e Ciências Exatas da Universidade Federal do Paraná, pela seguinte banca examinadora:

Orientador:

\_\_\_\_\_  
Prof. Dr. Anselmo Chaves Neto  
Departamento de Estatística, UFPR

Co-Orientador:

\_\_\_\_\_  
Prof. Dr. Alexandre Rasi Aoki  
PPGEE – UFPR/ Lactec

\_\_\_\_\_  
Prof<sup>a</sup>. Dr<sup>a</sup>. Thelma Solange Piazza Fernandes  
PPGEE – UFPR

\_\_\_\_\_  
Prof. Dr. Germano Lambert Torres  
PPGEE - UNIFEI

\_\_\_\_\_  
Prof. Dr. Reinaldo Castro Souza  
PPGEE – PUC-RIO

Curitiba, 12 de novembro de 2012

## DEDICATÓRIA

A Deus.

A minha Esposa

A minha Família.

Aos meus Amigos.

Dedico esse trabalho a minha querida esposa Elisângela, ao meu pai Orlando, a minha mãe Cleide e meu irmão

Woryk por estarem ao meu lado e por terem me apoiado nessa etapa da minha vida.

## AGRADECIMENTO

Agradeço a Deus pelo seu imenso amor e misericórdia revelado nas oportunidades que a vida me trouxe. Quero também agradecer:

A minha esposa Elisângela pelo carinho, compreensão e paciência durante essa etapa da minha vida

A minha família pelo incentivo, confiança e carinho.

Ao meu orientador o Prof. Dr. Anselmo Chaves Neto, pela orientação, incentivo e amizade desde o início do trabalho.

Ao meu co-orientador Dr. Alexandre Rasi Aoki, pela orientação e amizade que muito influenciou nesse trabalho.

A todos os professores do programa de Pós-Graduação em Métodos Numéricos em Engenharia, que ajudaram diretamente na realização desse trabalho.

Aos pesquisadores do LACTEC pela paciência e informação que foram fundamentais nesse trabalho em especial.

Aos amigos que sempre estiveram ao meu lado e que foram fundamentais para a realização desse trabalho com seu apoio e incentivo.

A COPEL Distribuição pela cessão de dados e ao LACTEC pelo suporte financeiro.

Por fim a todos que, de alguma forma, contribuíram com a realização dessa dissertação.

## EPÍGRAFE

*“O tempo se foi, a canção terminou.  
Pensei que tivesse algo mais a dizer...”*

Pink Floyd

## RESUMO

O presente trabalho tem por finalidade apresentar a proposta de uma arquitetura de Sistema Multiagentes (SMA) para o Reconhecimento e Classificação de Padrões aplicados em problemas de Engenharia de Potência. A arquitetura proposta está centrada na integração de agentes inteligentes capaz de gerenciar métodos de Reconhecimento e Classificação de Padrões. Para criação do Sistema Multiagentes para Reconhecimento e Classificação de Padrões foram definidos dois agentes independentes, capazes de melhorar o Reconhecimento e Classificação de Padrões individualmente, sendo que a interação destes dois agentes ao sistema possibilite uma melhora no processo global. O Agente Minerador de dados foi responsável pela filtragem dos dados e aplicação da Análise Fatorial como uma técnica de pré-processamento de informação. O Agente Decisor utiliza uma Lógica Nebulosa com o intuito de ponderar a classificação feita pelos Agentes Especialistas. Cada Agente Especialista é responsável por gerenciar uma técnica de Reconhecimento e Classificação de Padrões, obtendo assim as regras de classificação de uma nova informação, bem como a Taxa Real de Acerto que é obtida pelo método de Lachenbruch (*leave-one-out*). Para validar a eficácia de cada Agente Inteligente e do Sistema Multiagentes de Reconhecimento e Classificação de Padrões, foram utilizados dois bancos de dados que contemplam problemas de Reconhecimento e Classificação de Padrões na Engenharia de Sistemas de Potência. O primeiro banco de dados contempla os dados de uma campanha de medição realizada no estado do Paraná que contém informações sobre as características de demanda de transformadores, informações essas que são utilizadas na identificação e classificação dos transformadores que apresentam problemas de Distorções Harmônicas de Tensão. O segundo banco de dados contempla características sobre as estruturas metálicas de linhas de transmissão dentro do estado do Paraná, essas características são utilizadas como base de dados para analisar o risco de queda destas estruturas.

Palavras-chave: Reconhecimento e Classificação de Padrões, Estruturas Metálicas, Distorções Harmônicas de Tensão, Sistemas Multiagentes.

## **ABSTRACT**

This work presents the development of Multi Agents System (MAS) architecture for Recognition and Classification Patterns applied in problems on Power System. The architecture proposed is centered on integrating intelligent agents able to manage the methods of Recognition and Classification Patterns. For creating Multi Agent System for Recognition and Classification Patterns defined two independent agents, capable to improving the Recognition and Classification Patterns and that the interaction of these two agents to the system allows an improvement in the global process. The agent Data Mining was responsible for filtering the data and applying the Factor Analysis as a technique for pre-processing of information. The Decisor Agent uses a Fuzzy Logic in order to consider the classification made by the Agents Specialists. Each Agent Specialist is responsible for managing a technical Recognition and Classification Patterns, thus obtaining classification rules of the new information and the Real Rate of Correct which is obtained by the method of Lachenbruch (leave-one-out). To validate the effectiveness of each Intelligent Agent and Multi Agent System for Recognition and Classification Patterns, was used two databases that include problems Recognition and Classification Patterns in Power System. The first database includes data from a measurement campaign conducted in the state of Parana, which contains information on the characteristics of demand of transformers, which then are used in the identification and classification of transformers that have problems of Voltage Harmonic Distortions. The second database includes characteristics on the metallic structures of transmission lines in the state of Paraná, these characteristics are used as a database to analyze the risk of falling of these structures.

Key-words: Pattern Recognition and Classification, Metallic Structures, Voltage Harmonic Distortion, Multi Agents System.



## LISTA DE FIGURAS

FIGURA 2.1 – DESIGN DA METODOLOGIA DO AGENTE DO SISTEMA PEDA.....	26
FIGURA 2.2 – MODELO DE UM NEURÔNIO ARTIFICIAL .....	54
FIGURA 2.3 – FUNÇÕES DE ATIVAÇÃO DE UMA RNA .....	55
FIGURA 2.4 – REDE NEURAL ARTIFICIAL .....	57
FIGURA 2.5 – MODELO DE REDE NEURAL COM UMA CAMADA .....	57
FIGURA 2.6 – MODELO DE REDE NEURAL COM DUAS CAMADAS .....	58
FIGURA 2.7 – MODELO DE REDE NEURAL RECORRENTE .....	58
FIGURA 2.10 – CONJUNTO NEBULOSO DA VARIÁVEL “TEMPERATURA” .....	77
FIGURA 2.11 – CONFIGURAÇÃO BÁSICA DE UM CONTROLADOR NEBULOSO.	78
FIGURA 2.12 – DIAGRAMA TÍPICO DE UM MODELO DE INFERÊNCIA MAMDANI	79
FIGURA 3.1 – PERÍODO DE MONITORAÇÃO E PERÍODO DE ANÁLISE .....	85
FIGURA 3.2 – EXEMPLO DE DHTV – ABAIXO DE 6% .....	85
FIGURA 3.3 – EXEMPLO DE DHTV – ACIMA DE 6% .....	86
FIGURA 3.4 – ESPECTRO DA TENSÃO NO PERÍODO DE MAIOR DISTORÇÃO – ABAIXO DE 6%.....	87
FIGURA 3.5 – ESPECTRO DA TENSÃO NO PERÍODO DE MAIOR DISTORÇÃO – ACIMA DE 6%.....	87
FIGURA 3.6 – QUEDA DA ESTRUTURA 93 DA LT IRATI – SABARÁ.....	90
FIGURA 3.7 – QUEDA DA ESTRUTURA 34 DA LT MEDIANEIRA – CÉU AZUL .....	91
FIGURA 3.8 – IDENTIFICAÇÃO DAS LINHAS ONDE OCORRERAM QUEDAS.....	91
FIGURA 3.9 – IDENTIFICAÇÃO DAS LINHAS SELECIONADAS PARA ESTUDO... ..	92
FIGURA 3.10 – PROPOSTA DE ARQUITETURA DO SISTEMA MULTIAGENTES..	97
FIGURA 3.12 – DIAGRAMA DE INTERAÇÃO ENTRE OS AGENTES.....	102
FIGURA 3.13 – PROCESSO DO AGENTE DECISOR .....	103
FIGURA 3.14 – FUNÇÃO DE PERTINÊNCIA PARA TAXA REAL DE ACERTO.....	104
FIGURA 3.15 – FUNÇÃO DE PERTINÊNCIA PARA VARIÁVEL DE SAÍDA.....	105
FIGURA 3.16 – PROCESSO DOS AGENTES ESPECIALISTAS .....	107
FIGURA 4.1 – SELEÇÃO DO NÚMERO DE FATORES DOS DADOS DE TESTE .	113
FIGURA 4.2 – MÉTODO DE VALIDAÇÃO DA ARQUITETURA MULTIAGENTES .	117
FIGURA 4.3 – SELEÇÃO DO NÚMERO DE FATORES DOS DADOS DAS DHTV	122

## LISTA DE QUADROS

QUADRO 4.1 – CLASSIFICAÇÃO DOS DADOS ORIGINAIS PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL.....	111
QUADRO 4.2 – CLASSIFICAÇÃO DOS DADOS DAS ORIGINAIS PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL .....	111
QUADRO 4.3 – CLASSIFICAÇÃO DOS DADOS ORIGINAIS PELO MÉTODO DE REDES NEURAS COM TAXA DE ACERTO REAL .....	112
QUADRO 4.4 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL.....	114
QUADRO 4.5 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL.....	115
QUADRO 4.6 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAS COM TAXA DE ACERTO REAL.....	115
QUADRO 4.7 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE TESTE .....	115
QUADRO 4.8 – CLASSIFICAÇÃO DOS DADOS DE TESTE PELO AGENTE DECISOR COM TAXA DE ACERTO REAL .....	118
QUADRO 4.9 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE TESTE .....	118
QUADRO 4.10 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL.....	120
QUADRO 4.11 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL.....	120
QUADRO 4.12 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO PELO MÉTODO DE REDES NEURAS COM TAXA DE ACERTO REAL.....	120
QUADRO 4.13 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL .....	124
QUADRO 4.14 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL.....	124
QUADRO 4.15 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAS COM TAXA DE ACERTO REAL.....	125

QUADRO 4.16 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO.....	125
QUADRO 4.17 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL..	126
QUADRO 4.18 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL....	126
QUADRO 4.19 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL.....	126
QUADRO 4.20 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL .....	129
QUADRO 4.21 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL.....	129
QUADRO 4.22 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL.....	130
QUADRO 4.23 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS QUEDAS DE ESTRUTURAS	130
QUADRO 4.24 – RESUMO DA APLICAÇÃO DA ANÁLISE FATORIAL COMO AGENTE MINERADOR DE DADOS .....	130
QUADRO 4.25 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA LÓGICA NEBULOSA .....	132
QUADRO 4.26 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE DISTORÇÕES HARMÔNICAS DE TENSÃO.....	132
QUADRO 4.27 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL.....	133
QUADRO 4.28 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS QUEDAS DE ESTRUTURAS	133
QUADRO 4.29 – RESUMO DA APLICAÇÃO DA LÓGICA NEBULOSA COMO AGENTE DECISOR .....	134
QUADRO 4.30 – COMPARAÇÃO DOS RESULTADOS DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO.....	135
QUADRO 4.31 – COMPARAÇÃO DOS RESULTADOS DOS DADOS DAS QUEDAS DE ESTRUTURAS.....	135

## LISTA DE TABELAS

TABELA 2.1 – TABELA DE CLASSIFICAÇÃO .....	62
TABELA 3.1 – CARACTERÍSTICAS DA COPEL PARA A ATUAÇÃO NO PARANÁ. 81	
TABELA 3.2– CARACTERÍSTICAS DA COPEL – DADOS DA DISTRIBUIÇÃO .....	82
TABELA 3.3 – DISTRIBUIÇÃO DOS PONTOS DE MONITORAÇÃO .....	83
TABELA 3.4 – DISTRIBUIÇÃO DAS AMOSTRAS ENTRE ÁREAS URBANAS E RURAIS.....	84
TABELA 3.5 – VALORES MÁXIMOS, MÍNIMOS E P95% DA DHTV - DENTRO DO LIMITE (6%) .....	86
TABELA 3.6 – VALORES MÁXIMOS, MÍNIMOS E P95% DA DHTV - FORA DO LIMITE (6%) .....	86
TABELA 3.7 – NÚMERO DE TRANSFORMADORES URBANOS MEDIDOS.....	88
TABELA 3.8 – VARIÁVEIS ASSOCIADAS À CARACTERÍSTICA DO TRANSFORMADOR .....	89
TABELA 3.9 – VARIÁVEIS ASSOCIADAS À CARACTERÍSTICA DO CONSUMO... 89	
TABELA 3.10 – RELAÇÃO DE CARACTERÍSTICAS NÃO-NUMÉRICAS PARA NUMÉRICAS.....	93
TABELA 4.1 – BANCO DE DADOS PARA TESTE DA ARQUITETURA .....	110
TABELA 4.2 – ANÁLISE FATORIAL DOS DADOS DE TESTE .....	112
TABELA 4.3 – CARREGAMENTOS FATORIAIS ROTACIONADAOS DO BANCO DE DADOS DE TESTE .....	113
TABELA 4.4 – ANÁLISE FATORIAL DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO .....	121
TABELA 4.5 – CARREGAMENTOS FATORIAIS ROTACIONADAOS DAS DISTORÇÕES HARMONICAS DE TENSÃO.....	122
TABELA 4.6 – ANÁLISE FATORIAL DOS DADOS DAS QUEDAS DE ESTRUTURAS.....	127
TABELA 4.7 – CARREGAMENTOS FATORIAIS ROTACIONADAOS DAS QUEDAS DE ESTRUTURAS.....	128

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>16</b>
1.1 CASO I – DISTORÇÕES HARMÔNICAS DE TENSÃO .....	17
1.2 CASO II – QUEDAS DE ESTRUTURAS METÁLICAS DE LINHAS DE TRANSMISSÃO .....	18
1.3 OBJETIVOS .....	19
1.3.1 Objetivos Específicos .....	19
1.3 ESTRUTURA DA TESE .....	19
<b>2 REVISÃO DE LITERATURA</b> .....	<b>21</b>
2.1 SISTEMAS MULTIAGENTES .....	21
2.1.1 Conceitos: Terminologia e Definições .....	21
2.1.2 Agente Inteligente .....	23
2.1.3 Definição dos Sistemas Multiagentes.....	24
2.1.4 Implementação de Sistemas Multiagentes .....	24
2.1.5 Projetando Sistemas Multiagente .....	26
2.1.6 Estado da Arte Sobre Aplicações de Sistemas Multiagentes em Engenharia de Potência.....	27
2.2 RECONHECIMENTO DE PADRÕES.....	36
2.2.1 Análise Discriminante – Método de Fisher .....	36
2.2.1.1 Problema Geral da Classificação com Várias Populações .....	37
2.2.1.3 Análise Discriminante para Duas Populações – Método de Fischer .....	39
2.2.1.4 Discriminação Entre Diversas Populações .....	43
2.2.1.5 Escore Discriminante Quadrático de Classificação .....	46
2.2.2 Regressão Logística.....	48
2.2.2.1 Modelo de Regressão Logística.....	48
2.2.2.2 Modelo de Regressão Logística Simples .....	49
2.2.2.3 Modelo de Regressão Logístico Múltiplo.....	52
2.2.3 Redes Neurais Artificiais (RNAs) .....	53
2.2.3.2 Neurônio Artificial .....	54
2.2.3.3 Funções de Ativação de uma RNA .....	55
2.2.3.4 Características Gerais das RNAs .....	55
2.2.3.5 Treinamento de uma RNA de Múltiplas Camadas.....	59
2.2.3.6 Algoritmo <i>Back- Propagation</i> .....	59

2.3 AVALIAÇÃO DA FUNÇÃO DE CLASSIFICAÇÃO .....	60
2.3.1 Critério TPM ( <i>Total Probability of Misclassification</i> ).....	60
2.3.2 Abordagem de Lachenbruch.....	63
2.4 ANÁLISE FATORIAL.....	64
2.4.1 Modelo Fatorial Ortogonal.....	65
2.4.2 Método das componentes principais via matriz de covariância .....	68
2.4.3 Método das componentes principais via matriz de correlação amostral .....	69
2.4.4 Rotação ortogonal de fatores.....	71
2.4.4.1 Critério Varimax .....	71
2.4.4.2 Critério Quartimax .....	72
2.4.4.3 Critério orthomax .....	73
2.4.5 Determinação dos escores fatoriais .....	73
2.4.5.1 Método dos mínimos quadrados ponderados .....	74
2.5 LÓGICA NEBULOSA .....	75
2.5.1 Teoria dos conjuntos nebulosos.....	75
2.5.2 Variáveis linguísticas.....	76
2.5.3 Controle nebuloso.....	77
2.5.4 Modelo Linguístico do Tipo Mamdani.....	78
<b>3 MATERIAL E MÉTODO.....</b>	<b>81</b>
3.1 MATERIAL .....	81
3.1.1 Caso I – Campanha de Medição de Distorções Harmônicas de Tensão.....	81
3.1.1.1 Medições.....	82
3.1.1.2 Variáveis Utilizadas.....	88
3.1.2 Caso II – Histórico de Quedas de Estruturas de LTs no Estado do Paraná.....	90
3.1.2.1 Estruturas Metálicas em Operação na COPEL.....	92
3.1.2.2 Variáveis Utilizadas.....	92
3.2 MÉTODO.....	94
3.2.1 Introdução .....	94
3.2.2 Sistema Multiagentes.....	95
3.2.2.1 Arquitetura do Sistema Multiagentes Proposta .....	95
3.2.3 Agentes de Suporte.....	97
3.2.3.1 Agente Minerador de Dados.....	97
3.2.3.2 Agente de Comunicação .....	101
3.2.3.3 Agente Decisor .....	102

3.2.4 Agentes de Especialistas .....	107
3.2.4.1 Agente Especialista 1 – Regressão Logística .....	108
3.2.4.2 Agente Especialista 2 – Escore Quadrática.....	108
3.2.4.3 Agente Especialista – Redes Neurais.....	109
<b>4 RESULTADOS E DISCUSSÕES .....</b>	<b>110</b>
4.1 – TESTE DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES.....	111
4.2 VALIDAÇÃO DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES.....	119
4.2.1 Validação do Agente Minerador de Dados.....	119
4.2.1.1 Caso I – Distorções Harmônicas de Tensão .....	120
4.2.1.2 Caso II – Quedas de Estruturas Metálicas de Linhas de Transição .....	125
4.2.1 Validação do Decisor .....	131
4.2.1.1 – Caso I: Distorções Harmônicas de Tensão .....	131
4.2.1.2 - Caso II - Quedas de Estruturas Metálicas de Linhas de Transição.....	133
4.3 – RESULTADOS DA VALIDAÇÃO DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES .....	134
<b>5 CONCLUSÕES.....</b>	<b>137</b>
<b>6 REFERÊNCIAS BIBLIOGRÁFICAS.....</b>	<b>141</b>

## 1 INTRODUÇÃO

Durante as últimas décadas houve um considerável crescimento do interesse em problemas de Reconhecimento e Classificação de Padrões, com conseqüente aumento da necessidade de métodos para uso na elaboração de sistemas nesta área.

O Reconhecimento de Padrões é o processo de inverter a relação entre categoria geradora de padrão e o próprio padrão, ou seja, dado um vetor de medidas de dimensão  $p$  pergunta-se: qual a categoria entre várias possíveis gerou esse vetor? De forma que conhecendo a relação pode-se fazer a classificação de pessoas doentes, imagens, objetos, entre outros. O cérebro humano realiza esse processo instantaneamente, mas computacionalmente, esse processo não é tão rápido, nem tão preciso. Em virtude disso, muitas pesquisas foram e continuam sendo realizadas nessa área. O primeiro algoritmo de reconhecimento e classificação de padrões foi elaborado por Fisher, em 1936, e foi denominado de Função Discriminante de Fisher, o qual utiliza conceitos estatísticos.

Um método de reconhecimento de padrões deve se basear no conhecimento extraído dos itens pertencentes a uma base de dados classificados a priori e na classificação de um novo item, desconhecido até então. Dessa forma tem-se o desenvolvimento de métodos poderosos de reconhecimento de padrões, principalmente para tarefas como o reconhecimento de dígitos, reconhecimento de faces, falhas em equipamentos, tendências financeiras, entre outros.

Visto que em muitos problemas reais, apenas a utilização de um método de Reconhecimento e Classificação de Padrões, ou até mesmo comparação de vários métodos, não são suficientes para uma boa classificação torna-se importante o desenvolvimento de um sistema capaz de aplicar e gerenciar vários métodos de Reconhecimento e Classificação de Padrões. O presente trabalho tem como propósito principal apresentar e testar uma arquitetura de Sistema Multiagentes que permita combinar técnicas de Reconhecimento e Classificação de Padrões e Mineração de Dados, capaz de identificar as variáveis importantes para os modelos e sintetizar os resultados apresentados por cada técnica de Reconhecimento e Classificação de Padrões separadamente. Foram utilizados dados de dois problemas de Reconhecimento e Classificação de Padrões de Sistemas de Potencia que são apresentados a seguir.



## 1.1 CASO I – DISTORÇÕES HARMÔNICAS DE TENSÃO

Nos últimos anos, a Qualidade da Energia Elétrica (QEE) tem despertado o interesse dos órgãos reguladores do setor elétrico. Isso se deve pela forte e contínua degradação dos índices de conformidade e continuidade preconizados pelas normas vigentes. O desenvolvimento de equipamentos cada vez mais sensíveis às variações no fornecimento da energia elétrica, a crescente conscientização da população sobre seus direitos enquanto consumidores, aliados à escassez de documentos normativos que estabeleçam penalidades sobre violações de limites, são alguns dos fatores que contribuem para a mencionada preocupação. Igualmente, os reparos ou substituições de equipamentos elétricos, em face dos elevados níveis dos distúrbios, têm provocado custos financeiros indesejáveis. Os valores desses custos poderiam ser revertidos em investimentos para a melhoria do sistema elétrico da concessionária.

Dos fenômenos associados à Qualidade da Energia Elétrica, as distorções harmônicas de tensão merecem especial atenção, em virtude da crescente presença de cargas com características não-lineares nas redes elétricas de distribuição. Esta forte disseminação de cargas poluidoras, ou seja, cargas não-lineares expõem consumidores e concessionárias a níveis elevados de distorção harmônica, podendo provocar efeitos indesejáveis sobre o sistema, como o aumento das perdas elétricas, funcionamento inadequado de equipamentos, sobretensões harmônicas, dentre outros.

Diante do exposto, a quantificação e qualificação do conteúdo harmônico nas redes elétricas de distribuição tornam-se extremamente importantes para direcionar os esforços, na tentativa de mitigar ou reduzir o fenômeno, bem como, subsidiar os órgãos reguladores na especificação de limites aceitáveis. No entanto, para o cômputo dos patamares de distorção harmônica nas redes e instalações elétricas, via de regra, são realizadas séries de medições, intentando direcionar as ações para especificar soluções mitigadoras. Esta alternativa, contudo, nem sempre se mostra eficaz, dado a quantidade de pontos possíveis de monitorações em detrimento ao número de equipamentos disponíveis, sendo então importante a utilização de técnicas para a quantificação das distorções dentro do sistema.

## 1.2 CASO II - QUEDAS DE ESTRUTURAS METÁLICAS DE LINHAS DE TRANSMISSÃO

Um sistema de transmissão de energia elétrica compreende toda rede que interliga as usinas geradoras às subestações da rede de distribuição, sendo necessários milhares de quilômetros de linhas de transmissão (LT) que se interligam com diferentes níveis de tensão.

Os tipos de linhas de transmissão mais comuns são: linha aérea, em corrente alternada ou em corrente contínua com condutores separados pelo ar (meio dielétrico), e linha subterrânea, com um fio central condutor isolado de um condutor externo coaxial de retorno (FUCHS, 1977).

O sistema elétrico brasileiro é interligado através de linhas de transmissão aéreas por praticamente todo seu território. O estado do Paraná, por exemplo, possui mais de 7.500 km de linhas aéreas que se interligam por todo o estado transmitindo energia nas tensões de 69 kV, 138 kV, 230 kV, 345 kV e 525 kV.

Para que os cabos aéreos que transportam a energia sejam sustentados de forma eficaz e segura, são necessárias estruturas específicas para cada nível de tensão, para cada volume de carga a transportar e para diferentes tipos de relevos a serem transpostos. Assim, de acordo com as especificidades de cada projeto, as estruturas adotadas podem ser metálicas, de concreto ou mesmo de madeira, sendo definido de acordo com a vegetação regional e sua densidade demográfica.

Em um sistema de transmissão aéreo, existe uma série de riscos associados ao transporte da energia das fontes geradoras às fontes consumidoras finais. Dentre estes riscos, os principais são: desgastes em cabos e ferragens, colisões, descargas por aproximação excessiva de objetos aos condutores e quedas das torres aéreas devido a fatores climáticos, as quais acarretam descontinuidades de abastecimento e prejuízos de toda natureza. (WAZEN, 2008)

Normalmente as quedas de estruturas são identificadas após a ocorrência de fortes chuvas e vendavais. A queda das estruturas traz por consequência o rompimento dos cabos de energia ou, quando não há rompimento de cabos, faz com que ocorra um curto circuito fase-terra pelo contato de pelo menos um dos cabos com o corpo da estrutura ou mesmo com o solo.

Assim, tendo em conta o uso intensivo de linhas aéreas cujos riscos associados ao transporte podem levar ao desabastecimento de energia, uma análise

quantitativa dos mesmos com o intuito de identificá-los e mitigá-los tem muito a contribuir para o planejamento, operação e manutenção das linhas aéreas de transmissão de energia elétrica.

### 1.3 OBJETIVOS

O principal objetivo deste trabalho é apresentar uma arquitetura de um Sistema Multiagentes (SMA), que seja capaz de gerenciar métodos de Reconhecimento e Classificação de Padrões e Mineração de Dados das variáveis de entrada dos modelos bem como gerenciar a síntese dos resultados dos métodos aplicados.

#### 1.3.1 Objetivos Específicos

A fim de alcançar o objetivo geral estabeleceram-se os seguintes objetivos específicos:

1. Avaliar técnicas de Reconhecimento e Classificação de Padrões nos problemas de testes e validação;
2. Testar o modelo de Análise Fatorial para Mineração de Dados, bem como interpretar os resultados dos fatores explicativos do modelo;
3. Apresentar a arquitetura do Sistema Multiagentes para o Reconhecimento e Classificação de Padrões;
4. Testar a técnica de Lógica Nebulosa na síntese de classificação dos modelos de Reconhecimento e Classificação de Padrões.

#### 1.3 ESTRUTURA DA TESE

Os próximos capítulos estão estruturados da seguinte forma. O capítulo 2 apresenta a revisão de literatura que:

- Revisa os aspectos centrais dos Sistemas Multiagentes, bem como suas propriedades e seus conceitos. Neste capítulo também é apresentado um

estudo do atual “estado da arte” sobre as aplicações dos Sistemas Multiagentes na solução de problemas de Sistemas de Potência.

- Apresenta um estudo sobre os métodos de Reconhecimento e Classificação de Padrões aplicados no trabalho, bem como suas definições e propriedades.
- Apresenta um estudo sobre Análise Fatorial, metodologia aplicada como Mineração de Dados dos problemas de teste e validação.
- Apresenta um estudo sobre os conceitos fundamentais sobre Lógica Nebulosa, que constitui da técnica empregada na decisão do Sistema Multiagentes para Reconhecimento e Classificação de Padrões.

O capítulo 3 descreve a sistemática utilizada para a obtenção dos dados utilizados que serão utilizados para testes e validação da arquitetura do Sistema Multiagentes (SMA) de reconhecimento e classificação de padrões. Dois problemas reais foram considerados neste trabalho com o objetivo de testar e validar o sistema de reconhecimento e classificação de padrões apresentada.

Nesse capítulo também é abordada a construção da arquitetura do Sistema Multiagentes (SMA), bem como a função de cada agente do modelo.

O capítulo 4 apresenta os resultados da aplicação da técnica de Análise Fatorial como mineração de dados, os resultados dos métodos de Reconhecimento e Classificação de Padrões e os resultados do Sistema Multiagentes de Reconhecimento e Classificação de Padrões.

Finalmente no capítulo 5 são apresentadas as conclusões e finalmente as referências.

## 2 REVISÃO DE LITERATURA

### 2.1 SISTEMAS MULTIAGENTES

Por mais de uma década, a proposta de utilização dos Sistemas Multiagentes (SMA) na aplicação de problemas de engenharia de potência vem sendo relatada em artigos (*Institute of Electrical and Electronics Engineers – IEEE*). Essa tecnologia foi implementada em uma gama de aplicações, tais como: diagnóstico preliminar, condição de monitoração, restauração de sistemas de potência, simulações de mercado, controle de redes, automação, entre outros.

McArthur, et al (2007) apresentaram um trabalho com o intuito de investigar todas as questões acerca da aplicação de Sistemas Multiagentes em problemas de sistemas elétricos de potência. Esse trabalho contém três partes importantes para a compreensão dos Sistemas Multiagentes, as quais serão utilizadas para delinear esse tópico da revisão de literatura. São elas:

- Definição dos Sistemas Multiagentes e dos benefícios obtidos através da sua aplicação na engenharia de potência;
- Devido SMA ser uma técnica relativamente nova, apresenta uma investigação dos desafios técnicos que devem ser superados, com intuito da sua utilização de forma eficaz;
- A terceira parte apresenta recomendações e orientações sobre o uso adequado das normas, métodos de concepção e implementação dos Sistemas Multiagentes.

#### 2.1.1 Conceitos: Terminologia e Definições

Para se explorar os benefícios dos Sistemas Multiagentes na aplicação em engenharia de potência deve-se entender claramente os conceitos básicos sobre o sistema.

O fato de existir várias definições sobre o que seria um agente demonstra a dificuldade de definição. Mesmo existindo várias definições, todas elas compartilham de um conjunto básico de conceitos:

- Noção de um agente;
- Seu meio ambiente;
- Propriedade de autonomia.

Uma definição básica definida por Wooldridge e Weiss (1999) baseado nos trabalhos de Russell e Norvig (1995) e Maes (1995) revela que um agente é meramente “um *software* (ou *hardware*), entidade que está situado em algum ambiente e é autonomamente capaz de reagir às mudanças nesse ambiente”.

O ambiente é simplesmente tudo externo ao agente, ou seja, a fim de estar situado em um ambiente ou em pelo menos parte do ambiente a ser observável, ou de certo modo modificável. Esse ambiente pode ser físico (por exemplo, o sistema elétrico) e, portanto, observáveis através de sensores. Ou pode ser o ambiente interno como, por exemplo, fontes de dados, recursos computacionais e de outros agentes. Isso pode ser denominado de programa observável através de chamadas de sistema e mensagens.

Um agente pode alterar o ambiente, executando alguma ação: fisicamente, como fechar um ponto normalmente aberto de reconfiguração de uma rede, ou virtualmente como, o armazenamento de diagnóstico, informações em um banco de dados. O agente deve ter a capacidade de interagir com diferentes ambientes sem afetar a sua capacidade de raciocínio, nem os objetivos que foram projetados para alcançar. Isto significa que um agente pode operar em qualquer ambiente que suporta tarefas que ele tenha a intenção de executar.

Na definição de Wooldridge e Weiss (1999), uma entidade situada em um ambiente é apenas um agente que pode agir de forma autônoma, em resposta às mudanças ambientais. A autonomia é um termo um pouco evasivo, utilizado em todas as definições de agentes, mas raramente definido. Franklin e Graesser (1996) definem a autonomia de um agente como “exercícios de controle sobre suas próprias ações”, o que significa iniciar ações em horário determinado para a execução. Russell e Norvig (1995) vão mais longe, exigindo que a autonomia seja a resposta as ações de alguma mudança no ambiente e não apenas o resultado de algum conhecimento construído. Autonomia é, portanto, a capacidade de ação com base nas observações ambientais.

De uma perspectiva da engenharia, essa definição é problemática, pois não distingue claramente os agentes de software e sistemas de hardware já existentes. Indiscutivelmente, no âmbito da definição acima, alguns sistemas já existentes poderiam ser classificados como agentes, como, por exemplo, um relé de proteção pode ser considerado um agente, pois ele está situado no seu ambiente, ou seja, em um sistema de potência ele reage às mudanças no ambiente como alterações de tensão. Também apresenta um grau de autonomia. Argumentos semelhantes podem ser feitos por sistemas de software, tais como Unix e sistemas antivírus. Com isso há uma necessidade de saber como agentes e Sistemas Multiagentes diferem dos sistemas existentes, além disso, é preciso estudar quais as vantagens da aplicação da técnica em problemas na engenharia de potência.

### 2.1.2 Agente Inteligente

Com intuito de diferenciar os Sistemas Multiagentes com os já existentes, adota-se uma definição de agente inteligente proposta por Wooldridge e Weiss (1999), que estende o conceito de um agente definido acima para de um agente inteligente, alargando a definição de autonomia para autonomia flexível. Um agente inteligente, ou um agente que exhibe autonomia flexível, é composto pelas três características a seguir:

➤ Reatividade: um agente inteligente é capaz de reagir às mudanças no seu ambiente em tempo hábil e toma alguma ação com base nessas mudanças e da função a que se destina alcançar;

➤ Pró-atividade: um agente inteligente apresenta um objetivo direto de comportamento. O objetivo direto denota o comportamento de um agente que dinamicamente muda seu comportamento a fim de atingir os seus objetivos. Por exemplo, se um agente perde a comunicação com outro agente, cujos seus serviços sejam necessários para cumprir seus objetivos, ele irá procurar outro agente que fornece o mesmo serviço. Wooldridge e Weiss (1999) descrevem a pró-atividade como a capacidade do agente de “tomar a iniciativa”.

➤ Habilidade social: agentes inteligentes são capazes de interagir com outros agentes inteligentes. Capacidade social conota a simples passagem de dados entre diferentes *softwares* e entidades de *hardware*. Ela denota a capacidade de

negociar e interagir em uma maneira cooperativa. Essa capacidade é normalmente apoiada por uma linguagem de comunicação do agente, que permite que eles conversem entre si e não simplesmente transmitam dados.

Não só as características de reatividade, pró-atividade e habilidade social ajudam a distinguir os agentes dos sistemas tradicionais de *hardware* e *software*, como também a partir delas discuti-se nas seções a seguir que muitos de seus benefícios são derivados dessas qualidades.

### 2.1.3 Definição dos Sistemas Multiagentes

Os Sistemas Multiagentes compreendem de dois ou mais agentes ou agentes inteligentes. É importante reconhecer que não existe uma meta global do sistema, simplesmente as metas locais de cada agente separado. As intenções do projetista do sistema para o sistema só pode ser realizado através da inclusão de vários agentes inteligentes, com metas locais correspondentes a “pedaços”. Dependendo da definição específica de cada agente, os agentes em um Sistema Multiagentes podem ou não ter a capacidade de comunicar diretamente entre si. Porém, nas definições de agentes inteligentes proposta por Wooldridge e Weiss (1999), estes devem ter capacidade social e, portanto, devem ser capazes de se comunicar uns com os outros. Essa capacidade de interação entre agentes é o que diferencia claramente os Sistemas Multiagentes de outros tipos já existentes.

### 2.1.4 Implementação de Sistemas Multiagentes

Quando se trata de implementação dos Sistemas Multiagentes existem duas questões fundamentais a serem consideradas:

- Como deve ser construído um agente inteligente autônomo para aplicações na engenharia de potência?
- Como deve ser construída uma sociedade de agentes para aplicações na engenharia de potência?

Os agentes vêm sendo atualmente investigados por uma gama de aplicações dentro da comunidade de monitoramento e diagnóstico para o controle



de rede. A justificativa para a sua utilização muitas vezes encontra-se nas propriedades supostamente inerente à autonomia flexível, reatividade, pró-atividade, capacidade social, na natureza da distribuição de agentes, na possibilidade de comportamento emergente e na tolerância a falhas.

Porém, a escolha de um design específico para a aplicação da técnica pode restringir ao ponto de que essas propriedades não sejam exibidas. Por esta razão, é essencial que as melhores práticas atuais sejam seguidas ao desenvolver um Sistema Multiagentes. Assim, se faz necessária uma discussão sobre o atual estado da arte, abordando padrões de sistemas já existentes como o *Common Information Model* – CIM (2005), averiguando a melhor forma de permitir interoperabilidade entre agentes de diferentes designers e finalmente analisando a anatomia de um agente, que é identificado como uma área que requer mais pesquisas.

O uso de padrões se torna importante para o desenvolvimento dos Sistemas Multiagentes nas aplicações em engenharia de potência. Utilitários estão se esforçando para uma maior integração entre sistemas anteriormente separados. Normas recentes, como os sistemas de energia CIM (2005), promovem interfaces abertas entre sistemas de gestão de energia de diferentes fornecedores, e IEC 61850 (2005), promovem a interoperabilidade entre dispositivos dentro de subestações. Como a aplicação da tecnologia dos Sistemas Multiagentes está sendo amplamente difundida dentro da engenharia de potência, a adoção de normas que promovam a interoperabilidade entre os sistemas no futuro seria vantajoso, ou uma necessidade.

Nos últimos anos, a (FIPA) propõe normas para o desenvolvimento dos Sistemas Multiagentes diante da comunidade da ciência da computação. Em 2005, a FIPA foi formalmente aceita como comitê normativo da IEEE *Computer Society*.

A FIPA tem como objetivo definir as especificações e normas que podem ser utilizadas para apoiar a interoperabilidade entre os sistemas baseados em agente desenvolvido por diferentes empresas e organizações. Estas normas têm um grande impacto não somente sobre os métodos da comunicação entre agentes, mas também sobre como a arquitetura básica de um Sistema Multiagentes deve ser implementada.

### 2.1.5 Projetando Sistemas Multiagente

Desde meados de 1990, inúmeras metodologias surgiram para a especificação e design dos Sistemas Multiagentes. Metodologias de projeto dos Sistemas Multiagentes tendem a compartilhar algumas características comuns: uma fase de conceituação, onde o problema a ser resolvido é especificado; uma fase de análise; e uma fase de projeto que usa os resultados da fase de análise para a produção de detalhes de desenhos de diferentes agentes.

Na figura 2.1, a seguir temos a ilustração das diferentes fases da metodologia do projeto que McArthur, McDonald e Hossack,(2003) usou para especificar e projetar o sistema PEDA. Cada etapa da metodologia produz material que é usado nas fases posteriores do processo de criação. A metodologia se inicia com um conhecimento estruturado de engenharia, especificando os requisitos do sistema e capturando o conhecimento necessário para cumprir estes requisitos.

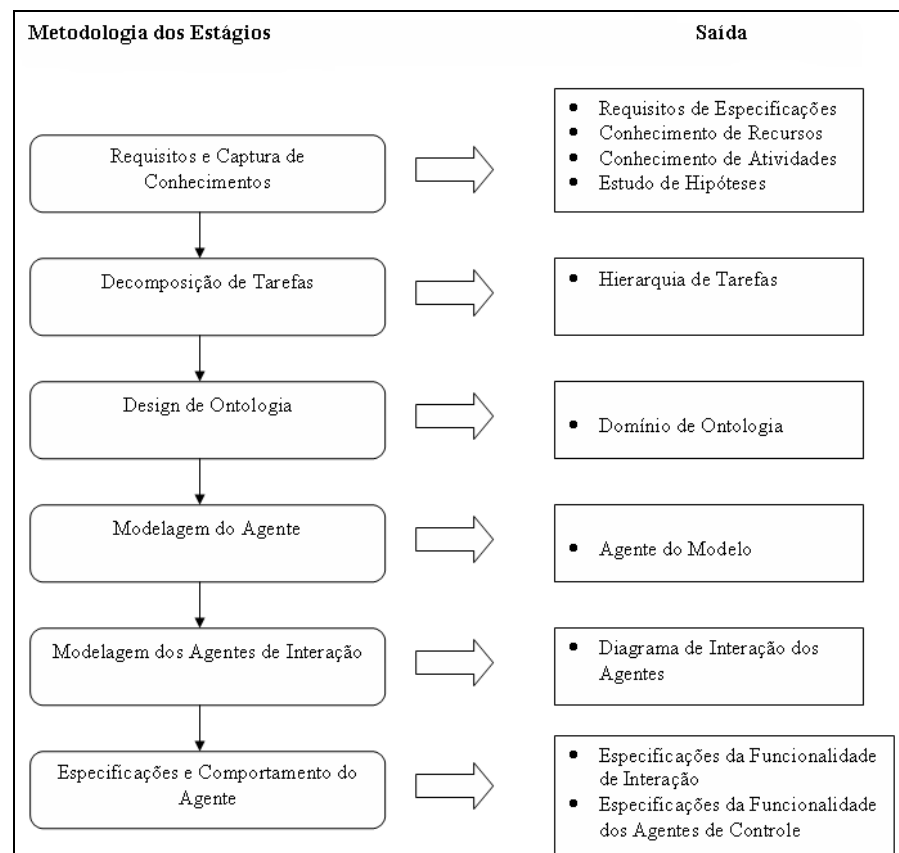


FIGURA 2.1 – DESIGN DA METODOLOGIA DO AGENTE DO SISTEMA PEDA  
 FONTE: MCARTHUR, MCDONALD E HOSSACK,(2003)

Durante o estágio de decomposição de tarefas, a especificação de requisitos e conhecimento adquirido na fase anterior são transformados em uma hierarquia de tarefas e subtarefas. Estas tarefas podem incluir as funções desempenhadas pelos sistemas legados. No caso do sistema PEDDA, o legado de sistemas inteligentes foi utilizado para fornecer dados de funções de análise. Após a decomposição de tarefas, a ontologia pode ser projetada.

A fase de modelagem do agente utiliza a hierarquia de tarefas e projeto de ontologia para identificar um grupo de agentes autônomos com as habilidades para executar as tarefas na hierarquia. Um agente pode encapsular uma ou mais tarefas e cada uma das tarefas na hierarquia deve ser atribuída a pelo menos um agente. O resultado é um conjunto de modelos de agentes que especifique as atribuições que os mesmos devem ser capazes de realizar. A metodologia também identifica as tarefas que podem ser atribuídas aos sistemas legados e para que os novos códigos precisem ser gerados.

Uma vez que os agentes tenham sido identificados, as interações dos agentes de apoio precisam ser definidas. A fase final do processo é a especificação da funcionalidade de interação do agente e da funcionalidade de controle do agente, isso equivale à especificação do comportamento que um agente deve exibir.

#### 2.1.6 Estado da Arte Sobre Aplicações de Sistemas Multiagentes em Engenharia de Potência

Para a investigação do atual “estado da arte” dos Sistemas Multiagentes nas aplicações na engenharia de potência foi realizado uma pesquisa utilizando como principais os anais da *Conferences Proceedings of the Intelligent Systems Application to Power Systems (ISAP)*, que é a principal conferência sobre Sistemas Multiagentes. E utilizados artigos das revistas *IEEE Transaction on Power Systems*, *IEEE Transactions on Power Delivery*, *IEEE Transaction on Energy Conversion* e *IEEE Transactions Evolutionary Computing*.

A partir dessa investigação foram detectadas quatro categorias de aplicações dos Sistemas Multiagentes na engenharia de potência. As quatro categorias são descritas a seguir:

- Monitoração e Diagnóstico

A principal área de aplicação dos Sistemas Multiagentes é a gestão e interpretação de dados nos problemas da engenharia de potência. A tecnologia dos Sistemas Multiagentes é uma excelente ferramenta para a coleta e a manipulação de informações e conhecimento.

O monitoramento de equipamentos do sistema nos oferece uma série de desafios, como por exemplo:

- Coleta de dados de uma variedade de sensores;
- Interpretação dos dados para extrair informações significativas.
- Combinar as informações de diferentes sensores para criar um diagnóstico geral;
- Fornecer as informações de diagnóstico em um formato correto e pertinente;
- Alterar automaticamente o sistema de energia.

Se considerarmos transformadores do sistema, existem vários sensores que podem ser utilizados para monitorá-los, como por exemplo: monitoração UHF de descargas parciais, monitoramento de descargas parciais, e gás dissolvido na medição de óleo. Além disso, informações operacionais sobre o carregamento do circuito e falhas também podem ser usados para informar o processo de diagnóstico. A tecnologia dos agentes permite a combinação de dados de todas estas fontes de forma flexível, ou seja, as informações são utilizadas quando a mesma estiver disponível e relevante, delegando a tarefa de acompanhar cada fonte um agente autônomo.

Como exemplo, um agente responsável pelo acompanhamento da saída dos sensores UHF informa o engenheiro ou ao algoritmo de diagnóstico quando uma atividade de descarga parcial significativa foi detectada. A autonomia do agente permite determinar quando tal informação deve ser comunicada, e para quem. A propriedade de flexibilidade permite a integração das informações de diagnóstico, e como essas informações podem ser utilizadas. Novos sensores e algoritmos de interpretação também podem ser introduzidos de modo homogêneo em todo o sistema, desde a arquitetura aberta permite extensibilidade.

Usando estes princípios, McArthur, Strachan e Jahn,(2004) desenvolveram um Sistema Multiagentes para monitoração de condições de transformador. Como uma idéia nova, os agentes de monitoramento de condição também podem ser capazes de modificar sua medida de tempo de processo, por exemplo, alterando a taxa de aquisição de dados. Enquanto a conexão física dos instrumento raramente podem ser modificados, num quadro de instrumentação virtual, o agente de monitoramento pode controlar a execução de instrumentos específicos. Isso traria muitas vantagens como, por exemplo, a otimização de recursos como bateria e sistemas de computação.

Lim, et al. (2007) propõe um Sistema Multiagentes para a coordenação de dispositivos de proteção de um sistema de distribuição de energia. Os agentes são incorporados em um dispositivo de proteção com o intuito de detectarem falhas ou de se adaptar na rede de uma forma autônoma, a fim de promover mudanças nos seus parâmetros de funcionamento para as novas condições de operação através da colaboração com outros agentes. Por exemplo, um agente associado a um relê de sobrecorrente, calcula a carga máxima de corrente quando uma mudança acontece na rede, e corrige os seus parâmetros.

Simulações de casos foram feitas para analisar os agentes e suas operações em um sistema de distribuição.

Häger, et al. (2009) apresenta uma aplicação dos Sistemas Multiagentes na coordenação de controladores de fluxos de potência. Cada agente é associado a dispositivos ao longo da rede que enviam mensagem contendo informações sobre o estado da topologia do sistema. Um agente de controle instalado em cada rede avalia as informações através de uma função de ponderação, a fim de apresentar uma análise sobre a topologia do sistema, que permite efetuar ações de controle em cada dispositivo da rede. A arquitetura utilizada combina as vantagens de uma resposta rápida ao sistema e um design totalmente distribuído, que é julgado como essencial para o controle do fluxo de potência, evitando assim ações errôneas após eventos graves ou situações inesperadas no sistema.

Baker, Catterson e McArthur, (2009) descrevem uma proposta de integração dos Sistemas Multiagentes na utilização de sensores sem fio para monitoramento. Este trabalho descreve os passos necessários para integrar uma arquitetura de rede sem fio com um sistema de monitoramento já existente.

Dois estudos de caso são apresentados para demonstrar diferentes cenários em que este sistema pode ser aplicado. O primeiro é um experimento de laboratório baseados periodicamente captura temperatura e os dados de vibração. O segundo estudo de caso apresenta uma proposta para uma rede de sensores sem fio que pode ser adicionada a uma já existente em uma subestação de transmissão no Reino Unido. Estes estudos de caso demonstram que uma abordagem dos Sistemas Multiagentes combinada com a rede de sensores sem fio pode ser usados para uma série de aplicações de monitoramento.

Catterson, et al. (2009) detalha um projeto de operação baseado em agentes para o acompanhamento de dois transformadores 275/132 Kv 180MVA situados no Reino Unido. A tecnologia do Sistema Multiagentes foi utilizada como suporte para integrar diferentes diagnósticos e módulos de detecção de anomalias, com cerca de 50 sensores de captura de dados a cada 5 minutos. Para apoiar o diagnóstico e a detecção de anomalias, um grupo de agentes responsável pelo tratamento de dados e detecção de anomalias foram integrados ao sistema.

O sistema foi implementado em uma subestação para obtenção de dados, interpretação de medidas on-line a cada 5 minutos.

Momoh, Alfred e Xia, (2009) apresentam o desenvolvimento de uma estrutura de um Sistema Multiagentes para detecção e diagnóstico de “Arcos falhas” nos sistema de potência a bordo de navios. Também é fornecido um gerenciamento de falhas através de agendamento, que mantém a operação do sistema de potência a bordo. Os regimes de controle do Sistema Multiagentes são baseados em técnicas de otimização baseada em regras.

O cenário para a implementação do protótipo foi desenvolvido e testado em um Sistema Integrado de Potência (*Power Integrated System - IPS*). O resultado, apresenta um aumento na eficiência da rede e a possibilidade minimizada de falhas não detectadas.

- Pós-diagnóstico de falhas do sistema de potência

A investigação sobre as causas e o impacto das falhas nos sistemas de potência apresentam um grande número de fontes de dados. Estas fontes de dados incluem a supervisão, controle e aquisição de dados, dados de gravador digital de falhas entre outros. De forma semelhante ao problema de monitoramento de

condições discutidas anteriormente, a automação da análise desses dados de apoio à decisão se torna essencial para o sistema. Davidson, et. al. (2006) apresenta um trabalho de um sistema do Reino Unido, que registrou um afluxo de 15 000 alarmes de supervisão e aquisição de dados e 1695 registros de falhas digital durante uma única tempestade. Os engenheiros requerem ferramentas de análise de apoio eficaz para combater tais situações.

A investigação sobre a aplicação de sistemas inteligentes para a análise dos dados de sistemas de potência está em curso por mais de duas décadas e tem produzido uma variedade de ferramentas e técnicas para analisar as fontes de dados individuais. A tecnologia do Sistema Multiagentes pode ser usada para integrar ferramentas de análise de dados, a fim de reforçar o apoio a diagnóstico, dando uma visão holística do desempenho de sistemas de energia com base em uma variedade de fontes de dados.

- Controle Distribuído

Com a introdução da geração de energia distribuída, controle de carga e operações de mercado a complexidade na distribuição de redes se torna cada vez maior. Os Sistemas Multiagentes fornece uma tecnologia flexível para o controle do sistema de potência moderno. A atual abordagem através de um sistema central SCADA e vários pequenos sistemas distribuídos já não são suficientes para o controle de determinadas operações. Com isso uma abordagem que proporciona um controle local adaptável, inteligente e rápido se torna necessária. As aplicações dos Sistemas Multiagentes aplicados no controle distribuído incluem:

- Restauração do sistema de alimentação;
- Operação de redes de distribuição de ativos;
- Microgrid e;
- Controle de sistemas elétricos de bordo.

Dimeas e Hatzargyriou, (2005) e Korbik, et al. (2005) apresentam uma abordagem baseada em agentes, como uma alternativa para a gestão centralizada de energia do sistema e controle. Com a gestão da distribuição e funcionalidade de controle utilizando agentes inteligentes, a tomada de decisão sobre restauração da

rede, reconfiguração, expedição de geração e gestão de cargas podem ser localmente gerenciados.

As tomadas de decisões locais exigem que os agentes sejam capazes de efetuar um conjunto de ações, tais como monitoração de condições locais, controle de aparelhos do sistema em coordenação com outras regiões da rede.

Davison e McArthur, (2007) descrevem a proposta de uma aplicação da tecnologia dos Sistemas Multiagentes em um sistema de gestão de uma rede regional autônoma AuRA-NMS. Esse sistema de gestão atualmente é desenvolvido numa parceria entre operadoras de distribuição, fabricantes de equipamentos e diversas universidades do Reino Unido. O uso da tecnologia dos Sistemas Multiagentes teve como objetivo fornecer um software flexível e extensível com o intuito de integrar os softwares de gerenciamento da rede AuRA-NMS.

Huang, et al. (2007) propõem uma arquitetura para os Sistemas Multiagentes aplicado na reconfiguração de um sistema baseado em malha estruturada do sistema de potência. Uma das características importantes da arquitetura proposta é que cada agente interage com um componente do sistema de potência e eles trabalham de uma forma completamente descentralizada, ou seja, não existe nenhum controle central do sistema. Nesta arquitetura se dois componentes tem ligação física entre si, então os agentes correspondentes são definidos como agentes vizinhos. Cada agente tem comunicação restrita com o seu agente vizinho.

Para testar a metodologia proposta uma plataforma de simulação de hardware foi desenvolvido. A plataforma de simulação baseia-se em simuladores digital em tempo real.

Dimeas e Hatziargyriou, (2007) apresentam uma proposta de utilização de um Sistema Multiagentes no controle de uma Usina Virtual (Virtual Power Plant). Esta Usina Virtual é formada por Micro-Grids e outras unidades de produção e consumo e tem como objetivo fazer com que varias unidades de produção que não estão conectadas cooperem como uma única unidade do sistema. O objetivo principal do trabalho apresenta as vantagens da utilização de agentes de controle da Usina Virtual. E apresenta através de exemplos e estudos de caso como a inteligência local e a capacidade social dos agentes podem fornecer soluções para o controle ótimo e eficaz da Usina Virtual.



Rosa, et al. (2009) apresentam uma proposta de construção de um ambiente baseado em um Sistema Multiagentes, a fim de melhorar e clarificar as discussões sobre o agendamento de manutenção, prevenção no sistema de potência. As decisões de manutenções do sistema em geral apresentam conflitos entre a empresa de geração e o Operador do Sistema de Transmissão. A empresa de geração, por exemplo, é responsável pela sua manutenção, mas é tarefa do Operador do Sistema de Transmissão agendar o período mais adequado para a manutenção. Esta relação é por vezes conflitante e pode ser visto como um obstáculo à maximização dos lucros das empresas de geração, bem como ao Operador do Sistema de Transmissão, que é responsável por manter a confiabilidade do sistema global.

A fim de propor uma maneira de melhorar esse relacionamento, o trabalho propõe a criação de um ambiente baseado em um Sistema Multiagentes, onde as Empresas Geradoras e o Operador do Sistema de Transmissão podem interagir afim de atingir um cronograma mais adequado para manutenções.

Lim, et. al. (2009) apresentam um método para melhorar a eficiência da restauração distribuída inteligente de um Sistema de Distribuição Automatizado (DAS) com base num Sistema Multiagentes. Restauração distribuída inteligente é um novo método de restauração que utilizam mútuas troca de informações, seção de separação automática de falhas, com o objetivo de apresentar uma restauração mais rápida e confiável.

Lo, Wang e Lu, (2009) apresentam um Sistema Multiagentes aplicado na restauração de falha com e sem a assistência de geração distribuída e compara seu desempenho com um sistema de processamento centralizado. Os resultados de simulação indicam que, para detecção de falhas, isolamento e recuperação de serviço um Sistema Multiagentes pode superar o processamento centralizado do sistema.

Os resultados da simulação mostram que o Sistema Multiagentes foi mais eficiente que o sistema de processamento centralizado. Este estudo mostra que, quando equipado com equipamentos de sincronização adequada, a geração distribuída pode ser ativada para a operação de ilhamento intencional e/ou coordenado com o apoio de alimentadores para reduzir a carga (un-served) durante as falhas.

Du, et al. (2009) apresentam os resultados do estudo da viabilidade de um novo conceito de colaboração on-line para controle de estabilidade de tensão no sistema de potência. Nesse estudo os autores propõe a implementação dos Sistemas Multiagentes usando a teoria de grafos para a execução de controle do sistema de potência e a sua comunicação em tempo real. Os resultados dos testes on-line colaborativa proposta de controle de estabilidade de tensão no caso das 10 máquinas do sistema de alimentação de 39 nós da Nova Inglaterra são apresentados. Resultados de um estudo de viabilidade por meio de uma simulação são apresentados que levam em consideração as falhas em cascata de baixa probabilidade de o sistema de energia.

- Modelagem e Simulação

Dentro dos sistemas de potência atuais a modelagem e a simulação de situações reais se torna muito complexa com a utilização de métodos tradicionais. Por esta razão, o uso dos Sistemas Multiagentes na modelagem e nas simulações se torna um grande benéfico para modelagem e simulações de sistemas de potência. Estas aplicações têm uma propriedade comum: o comportamento global do sistema é muito complexo, mas é gerada pela interação de entidades mais simples.

Widergren, et al. (2004) apresenta uma abordagem para modelagem aplicada em simulações no mercado de energia, onde os agentes são vistos como fornecedores, corretores, geradores e os cliente.

Recentemente, a tecnologia dos Sistemas Multiagentes tem sido sugerida por Hopkinson, et al. (2006), Vlachogiannis, Hatziaargyiou e Lee, (2005), Koritarov, (2004) e McArthur, et al. (2003) para a integração e coordenação de diferentes modelos e modelagem de software.

Gehrke e Bindner, (2007) apresentam uma arquitetura de um Sistema Multiagentes em uma instalação experimental projetada para pesquisa sobre a tecnologia de controle de sistemas distribuídos de potência (SYSLAB). A instalação do SYSLAB é construída em torno de uma plataforma de controle descentralizado, a fim de permitir a análise de arquiteturas de controle diferentes, e usa a tecnologia dos agentes no papel de um corretor de serviço.

Van Sickel e Lee, (2009) apresentam o desenvolvimento de varias arquitetura de um Sistema Multiagentes baseados em *Distributed Real-time Agent Framework with Time Warp -DRAFT* . O sistema é baseado em abordagens de design comum tomadas no desenvolvimento de complexos sistemas distribuídos em tempo real. A simulação de eventos discretos permite uma rápida e simulação e sincronização, garantindo simultaneamente que todas as tarefas sejam cumpridas. Os agentes DRAFT podem passar da operação padrão e operar em tempo pseudo-real ou em tempo real para desenvolver todos os aspectos de um Sistema Multiagentes em tempo real. Os agentes DRAFT também apresentam a capacidade de simulação de agentes independentes que poderiam ser usados para melhorar a funcionalidade global do Sistema.

O uso de agentes DRAFT mostrou ser uma via eficaz para combinar a simulação contínua de uma usina com a natureza discreta de um Sistema Multiagentes.

- Proteção

A interação entre os Sistemas Multiagentes e dispositivos para proteção dos sistemas de potência é explorada nos documentos de Park e Lim (2006), Sheng et al. (2006), Coury et al. (2002), Fukui, et al. (1998) e Giovanini, et al. (2006). Em todos os documentos citados acima, relés de proteção e equipamentos associados são tratados como agentes e sua funcionalidade adicionadas de acordo com suas necessidades. Nesses estudos está sendo investigando a tecnologia de Sistemas Multiagentes como uma ferramenta para desenvolver esquemas de proteção tolerantes a falhas e auto-coordenados.

Edwards, et al. (2007), descrevem uma arquitetura para um software de segurança multiníveis baseado em agentes. O objetivo dessa arquitetura é evitar ataques conhecidos e reduzir ou eliminar consequência de ataques bem sucedidos, a fim de tornar a infra-estrutura elétrica mais segura sem comprometer as vantagens da sua produtividade. Para validação do modelo de segurança um protocolo experimental é apresentado para validar a arquitetura do modelo de segurança. Os resultados do teste mostram a eficácia da arquitetura desenvolvida.

## 2.2 RECONHECIMENTO DE PADRÕES

Para MULLER (1997), o reconhecimento de padrões é a área de pesquisa que tem por alvo identificar e discriminar objetos (padrões) em certo número de categorias ou classes. Algumas das aplicações são:

- Reconhecimento de padrões em imagens digitais;
- Análise do desempenho empresarial;
- Concessão de crédito;
- Diagnóstico médico preliminar;
- Elaboração de perfis de consumidores;
- Processamento de sinais;
- Reconhecimento de fala e de rosto;
- Inspeção visual para automação industrial.

Em geral, deseja-se que um problema de reconhecimento de padrões bem definido e restrito possua pequenas variações dentro da mesma categoria e grandes variações entre as diferentes categorias.

Segundo Chaves Neto (1997), estatisticamente, a construção de regras de reconhecimento e classificação de padrões para duas populações podem ser baseadas em quatro métodos principais: Análise Discriminante – Método de Fischer, Escore Discriminante Quadrático, Regressão Logística e Método das k-médias. Podem ser citadas também as técnicas de Inteligência Artificial para reconhecimento e classificação de padrões como, Redes Neurais, Conjuntos Aproximados (*Rough Sets*), entre outras.

### 2.2.1 Análise Discriminante – Método de Fisher

Segundo Johnson e Wichern (1998), a Análise Discriminante é uma técnica multivariada que tem por objetivo tratar dos problemas relacionados em separar conjuntos distintos de objetos (itens ou observações) e alocar novos objetos em conjuntos previamente definidos. Quando empregada como procedimento de classificação não é uma técnica exploratória, uma vez que ela conduz a regras bem definidas, as quais podem ser utilizadas para classificação de outros objetos.

Tem como objetivos imediatos, quando usada para discriminação e classificação, os seguintes:

1. Descrever algebricamente ou graficamente as características diferenciais dos objetos (observações) de várias populações conhecidas a fim de achar “discriminantes” cujos valores numéricos sejam tais que as populações possam ser separadas tanto quanto possível.

2. Agrupar os objetos (observações) dentro de duas ou mais classes determinadas. Tenta-se encontrar uma regra que possa ser usada na alocação ótima de um novo objeto (observação) nas classes consideradas. Uma função que separa pode servir para alocar um objeto e, da mesma forma, uma regra alocadora pode sugerir um procedimento discriminatório.

Na prática, os objetivos 1 e 2, freqüentemente, sobrepõem-se e a distinção entre separação e alocação torna-se confusa.

A Análise Discriminante foi originalmente desenvolvida na Botânica tendo por objetivo fazer a distinção de grupos de plantas com base no tamanho e no tipo de folha o que tornaria possível, posteriormente, classificarem as novas espécies encontradas.

### 2.2.1.1 Problema Geral da Classificação com Várias Populações

Considere:

- $f_i(\underline{X})$  a função densidade de probabilidade associada à população  $\Pi_i$  com  $i = 1, \dots, g$ ;
- $p_i$  a probabilidade a priori da observação pertencer à população  $\Pi_i$ ;
- $R_k$  o conjunto dos  $\underline{X}$  classificados em  $\Pi_k$ ;
- $c(k | i)$  o custo de classificação de uma observação em  $\Pi_k$  quando na verdade pertence a  $\Pi_i$ , com  $i = 1, \dots, g$  e  $c(i | i) = 0$  e
- $P(k | i)$  a probabilidade de se classificar uma observação de  $\Pi_i$  em  $\Pi_k$  sendo:

$$P(k | i) = \int_{R_k} f_i(\underline{X}) dx \quad i = 1, \dots, g \quad (2.1)$$

O Custo Esperado de Reconhecimento Errado (*ECM*) ao se classificar  $\underline{X}$  de  $\Pi_i$  em  $\Pi_k$ , com  $k = 1, \dots, g$  é dado por:

$$ECM(1) = P(2|1)c(2|1) + \dots + P(g|1)c(g|1) = \sum_{k=2}^g P(k|1)c(k|1) \quad (2.2)$$

De maneira análoga, obtem-se o  $ECM(2)$ ,  $ECM(3)$ , ...,  $ECM(g)$ . Considerando que  $ECM(k)$  ocorre com probabilidade  $p_k$  com  $k = 1, 2, \dots, g$  então:

$$ECM = p_1 ECM(1) + p_2 ECM(2) + \dots + p_g ECM(g) \quad (2.3)$$

logo,

$$P(k|i) = \int_{R_k} f_i(\underline{X}) dx \quad i = 1, \dots, g \quad (2.4)$$

$$ECM = p_1 \left( \sum_{k=2}^g P(k|1)c(k|1) \right) + \sum_{\substack{k=1 \\ k \neq 2}}^g P(k|2)c(k|2) + \dots + \sum_{k=1}^{g-1} P(k|g)c(k|g) \quad (2.5)$$

A regra de classificação consiste em se determinar  $R_k$  (com  $k = 1, 2, \dots, g$ ) que tornem o  $ECM$  mínimo tanto quanto possível, isto significa alocar  $\underline{X}$  na população  $\Pi_k$  de forma que:

$$\sum_{\substack{k=1 \\ k \neq 2}}^g P_i f(\underline{X}) c(k|i) \quad (2.6)$$

é menor e, num caso de empate,  $\underline{X}$  pode ser classificado em qualquer uma das duas populações para as quais o empate ocorre.

No caso de todos os custos serem iguais, é suficiente alocar  $\underline{X}$  na população  $\Pi_k$  em que:

$$\sum_{\substack{i=1 \\ i \neq k}}^g P_i f(\underline{X}) \quad (2.7)$$

é menor. Mas essa quantidade será menor quando o termo excluído  $P_i f_i(\underline{X})$  for maior. Assim, a regra de classificação do ECM mínimo com custos iguais por falhas de classificação é alocar  $\underline{X}$  em  $\Pi_k$  se:

$$P_k f_k(\underline{X}) > P_i f_i(\underline{X}) \text{ para todo } i \neq k \quad (2.8)$$

ou, alternativamente, alocar  $\underline{X}$  em  $\Pi_k$  se:

$$\ln P_k f_k(\underline{X}) > \ln P_i f_i(\underline{X}) \text{ para todo } i \neq k \quad (2.9)$$

### 2.2.1.3 Análise Discriminante para Duas Populações – Método de Fischer

A idéia de Fischer foi transformar observações multivariadas  $\underline{X}'_s$  em observações univariadas  $Y'_s$  tais que os  $Y'_s$  das populações  $\Pi_1$  e  $\Pi_2$  sejam separadas tanto quanto possível. Fischer tomou combinações lineares de  $\underline{X}$  para criar os  $Y'_s$ , dado que as combinações lineares são funções de  $\underline{X}$  e por outro lado são de fácil cálculo. Assim sendo  $\mu_{1y}$  a média dos  $Y'_s$  obtidos dos  $\underline{x}'_s$  pertencentes a  $\Pi_1$  (população 1) e  $\mu_{2y}$  a média dos  $Y'_s$  obtidos dos  $\underline{x}'_s$  pertencentes a  $\Pi_2$  (população 2), Fischer selecionou a combinação linear que maximiza a distância quadrática entre  $\mu_{1y}$  e  $\mu_{2y}$  relativamente à variabilidade dos  $Y'_s$ . Assim, seja:

$$\mu_1 = E(\underline{X}|\Pi_1) = \text{valor esperado em uma observação multivariada em } \Pi_1. \quad (2.10)$$

$$\mu_2 = E(\underline{X}|\Pi_2) = \text{valor esperado de uma observação multivariada em } \Pi_2. \quad (2.11)$$

e supondo a matriz de covariância  $\Sigma = E[(\underline{X} - \underline{\mu}_i)(\underline{X} - \underline{\mu}_i)']$   $i = 1, 2$  como sendo a mesma para as duas populações, e ainda considerando a combinação linear:

$$Y_{1 \times 1} = \underline{c}'_{1 \times p} \underline{X}_{p \times 1} \quad (2.12)$$

tem-se :

$$\mu_{1y} = E(Y|\Pi_1) = E(\underline{c}' \underline{X}|\Pi_1) = \underline{c}' E(\underline{X}|\Pi_1) = \underline{c}' \underline{\mu}_1 \quad (2.13)$$

$$\mu_{2y} = E(Y|\Pi_2) = E(\underline{c}' \underline{X}|\Pi_2) = \underline{c}' E(\underline{X}|\Pi_2) = \underline{c}' \underline{\mu}_2 \quad (2.14)$$

e

$$V(Y) = \sigma_y^2 = V(\underline{c}' \underline{X}) = \underline{c}' V(\underline{X}) \underline{c} = \underline{c}' \Sigma \underline{c} \quad (2.15)$$

que é supostamente a mesma para ambas as populações. Então, segundo Fischer, a melhor combinação linear é a derivada da razão entre o “quadrado da distância entre as médias” e a “variância de Y”.

$$\frac{(\mu_{1y} - \mu_{2y})^2}{\sigma_y^2} = \frac{(\underline{c}' \underline{\mu}_1 - \underline{c}' \underline{\mu}_2)^2}{\underline{c}' \Sigma \underline{c}} = \frac{\underline{c}' (\underline{\mu}_1 - \underline{\mu}_2) (\underline{\mu}_1 - \underline{\mu}_2)' \underline{c}}{\underline{c}' \Sigma \underline{c}} = \frac{(\underline{c}' \underline{\delta})^2}{\underline{c}' \Sigma \underline{c}} \quad (2.16)$$

onde:

$$\underline{\delta} = \underline{\mu}_1 - \underline{\mu}_2.$$

Assim, como  $\underline{\delta} = \underline{\mu}_1 - \underline{\mu}_2$  e  $Y = \underline{c}' \underline{X}$  tem-se que  $\frac{(\underline{c}' \underline{\delta})^2}{\underline{c}' \Sigma \underline{c}}$  é maximizada por:

$$\underline{c} = k \Sigma^{-1} \underline{\delta} = k \Sigma^{-1} (\underline{\mu}_1 - \underline{\mu}_2) \quad \text{para qualquer } k \neq 0. \quad (2.17)$$

Logo por multiplicidade com  $k = 1$  tem-se:

$$\underline{c} = \Sigma^{-1} (\underline{\mu}_1 - \underline{\mu}_2) \quad \text{e} \quad Y = \underline{c}' \underline{X} = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{X}, \quad (2.18)$$

que é conhecida como Função Discriminante Linear de Fisher. Ela transforma as populações multivariadas  $\Pi_1$  e  $\Pi_2$  em populações univariadas, tais que as médias destas populações são separadas tanto quanto possível relativamente à variância populacional, considerada comum. E, para classificar a observação multivariada  $\underline{x}_0$  usa-se o modelo:

$$y_0 = (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{x}_0 \quad (2.19)$$



onde  $y_0$  é o valor da Função Discriminante de Fisher para a nova observação  $\underline{x}_0$ , e considerando-se o ponto médio entre as médias das duas populações univariadas,

$$m = \frac{1}{2}(\mu_{1y} + \mu_{2y}), \quad (2.20)$$

como

$$m = \frac{1}{2}(\underline{c}'_1 \underline{\mu}_1 + \underline{c}'_2 \underline{\mu}_2), \quad (2.21)$$

$$m = \frac{1}{2}[(\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{\mu}_1 + (\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} \underline{\mu}_2] \quad (2.22)$$

$$m = \frac{1}{2}[(\underline{\mu}_1 - \underline{\mu}_2)' \Sigma^{-1} (\underline{\mu}_1 + \underline{\mu}_2)] \quad (2.23)$$

tem-se que:

$$E(y_0 | \Pi_1) - m \geq 0 \quad (2.24)$$

e

$$E(y_0 | \Pi_2) - m < 0 \quad (2.25)$$

ou seja, se  $\underline{x}_0$  pertence a  $\Pi_1$  espera-se que  $y_0$  seja igual ou maior do que o ponto médio  $m$ . Por outro lado se  $\underline{x}_0$  pertence a  $\Pi_2$ , o valor esperado de  $y_0$  será menor que o ponto médio. Portanto, a regra de classificação é:

- alocar  $\underline{x}_0$  em  $\Pi_1$  se  $y_0 - m \geq 0$ ;
- alocar  $\underline{x}_0$  em  $\Pi_2$  se  $y_0 - m < 0$ .

Os parâmetros  $\underline{\mu}_1$ ,  $\underline{\mu}_2$  e  $\Sigma$  geralmente são desconhecidos. Então, supondo que se tem  $n_1$  observações da variável aleatória multivariada  $\underline{x}_1$  de dimensão  $p$ , ou seja, tem-se uma amostra aleatória da população  $\Pi_1$  e  $n_2$  observações da variável aleatória multivariada  $\underline{x}_2$  de dimensão  $p$  que corresponde a uma amostra aleatória de população  $\Pi_2$ , os resultados amostrais correspondentes são:

$$\bar{\underline{x}}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} \underline{x}_{i1}; \quad S_1 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (\underline{x}_{i1} - \bar{\underline{x}}_1)(\underline{x}_{i1} - \bar{\underline{x}}_1)' \quad (2.26)$$

$$\bar{\underline{x}}_2 = \frac{1}{n_2} \sum_{i=1}^{n_2} \underline{x}_{i2}; \quad S_2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (\underline{x}_{i2} - \bar{\underline{x}}_2)(\underline{x}_{i2} - \bar{\underline{x}}_2)' \quad (2.27)$$

Assumindo que as populações sejam assemelhadas, é natural considerar a variância como a mesma, daí estima-se a matriz de covariância comum  $\Sigma$  pela matriz de covariância amostral calculada com as amostras conjuntas,

$$S_p = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2}{(n_1 + n_2 - 2)} \quad (2.28)$$

que é um estimador não –viciado daquele parâmetro  $\Sigma$ .

Consequentemente, a Função Discriminante Linear de Fischer Amostral é dada por:

$$\hat{Y} = \hat{c}\underline{X} = (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' S_p^{-1} \underline{x} \quad (2.29)$$

e a estimativa do ponto médio entre as duas médias amostrais univariadas,

$$\bar{y}_1 = \hat{c}\bar{\underline{x}}_1 \quad (2.30)$$

e

$$\bar{y}_2 = \hat{c}\bar{\underline{x}}_2 \quad (2.31)$$

é dada por:

$$\begin{aligned} \hat{m} &= \frac{1}{2}(\bar{y}_1 + \bar{y}_2) = \frac{1}{2} [(\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' S_p^{-1} \bar{\underline{x}}_1 + (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' S_p^{-1} \bar{\underline{x}}_2] \\ &= \frac{1}{2} (\bar{\underline{x}}_1 - \bar{\underline{x}}_2)' S_p^{-1} (\bar{\underline{x}}_1 + \bar{\underline{x}}_2) \end{aligned} \quad (2.32)$$

Finalizando, a regra de classificação do item caracterizado pelo vetor  $\underline{x}_0$  é:

Se  $\hat{y}_0 - \hat{m} \geq 0$  então  $\underline{x}_0$  é alocado em  $\Pi_1$

ou

Se  $\hat{y}_0 - \hat{m} < 0$  então  $\underline{x}_0$  é alocado em  $\Pi_2$

A combinação linear particular  $Y = \hat{c}' \underline{x} = (\bar{x}_1 - \bar{x}_2)' S_p^{-1} \underline{x}$  maximiza a razão:

$$\frac{(\bar{y}_1 - \bar{y}_2)^2}{S_y^2} = \frac{(\hat{c}_1 \bar{x}_1 - \hat{c}_2 \bar{x}_2)}{\hat{c}' S_p \hat{c}} = \frac{(\hat{c}' \underline{d})^2}{\hat{c}' S_p \hat{c}} \quad (2.33)$$

onde:

$$\underline{d} = \bar{x}_1 - \bar{x}_2 \quad (2.34)$$

e

$$S_y^2 = \frac{\sum_{i=1}^{n_1} (y_{i1} - \bar{y}_1)^2 + \sum_{i=1}^{n_2} (y_{i2} - \bar{y}_2)^2}{n_1 + n_2 - 2} \quad (2.35)$$

#### 2.2.1.4 Discriminação Entre Diversas Populações

O método anterior que é válido para duas populações  $g=2$ , pode ser estendido para diversas populações  $g \geq 2$ . O primeiro objetivo de Fischer com o método foi o de separar populações, podendo ser usado também para classificar novos itens em uma das populações. Esse método não necessita da suposição de que as diversas populações sejam normais multivariadas, porém é necessário assumir que as matrizes de covariância populacionais sejam iguais, ou seja,  $\Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$ .

Assim é preciso testar a hipótese de igualdade das matrizes de covariância, ou seja  $H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$ . Então tem-se o seguinte teste:

Hipótese nula:  $H_0 : \Sigma_1 = \Sigma_2 = \dots = \Sigma_g = \Sigma$

Hipótese alternativa: Pelo menos uma das matrizes é diferente das demais.

A estatística de teste é:

$$\text{Estatística de Teste } M = \left( \frac{|S_i|}{|S_p|} \right)^{\frac{(n_i-1)}{2}}$$

onde:

- $S_i$  é a estimativa da matriz de covariância do grupo  $i$ ;
- $S_p$  é a estimativa da matriz de covariância conjunta;
- $n_i$  é o número de observações do grupo  $i$ . (tamanho da amostra)

A transformação logarítmica de  $M$  conduz a estatística abaixo que é de tratamento matemático mais adequado:

$$B = (1 - c) \left\{ \left[ \sum_{i=1}^g (n_i - 1) \ln |S_p| - \sum_{i=1}^g [(n_i - 1) \ln |S_i|] \right] \right\} \sim \chi_{\frac{1}{2}p(p+1)(g+1)}^2 \quad (2.36)$$

com

$$c = \left[ \sum_{i=1}^g \frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^g (n_i - 1)} \right] \left[ \frac{2p^2 + 3p - 1}{6(p+1)(g-1)} \right] \quad (2.37)$$

onde:

- $p$  é a dimensão do vetor  $\underline{x}$  (número de variáveis);
- $n_i$  é o tamanho do grupo  $i$ ;
- $g$  é o número de grupos.

Então uma vez aceita a hipótese nula  $H_0$  de igualdade entre as matrizes de covariância, seja  $\underline{\bar{\mu}}$  o vetor médio dos diversos grupos (populações), representado por:

$$\underline{\bar{\mu}} = \frac{1}{g} \sum_{i=1}^g \underline{\mu}_i \quad (2.38)$$

e  $B_0$  a matriz “Soma de produtos cruzados entre grupos populacionais” tal que:

$$B_0 = \sum_{i=1}^g (\underline{\mu}_i - \underline{\bar{\mu}})(\underline{\mu}_i - \underline{\bar{\mu}})' \quad (2.39)$$

a combinação linear  $Y = \underline{c}' \underline{x}$  tem por esperança:

$$E(Y) = \underline{c}' E(\underline{x} | \Pi_i) = \underline{c}' \underline{\mu}_i \quad (2.40)$$

para população  $\Pi_1$  e variância:

$$V(Y) = \sigma_y^2 = \underline{c}'V(\underline{X})\underline{c} = \underline{c}'\Sigma\underline{c} \quad (2.41)$$

para todas as populações. Desta forma, o valor esperado  $\mu_{iy} = \underline{c}'\underline{\mu}_i$  muda quando a população da qual  $\underline{X}$  selecionado é outra. Tem-se então uma média global:

$$\bar{\mu}_y = \frac{1}{g} \sum_{i=1}^g \mu_{iy} = \underline{c}'\bar{\underline{\mu}} \quad (2.42)$$

e a razão entre a “Soma dos quadrados das distâncias das populações para a média global” e a variância de  $Y$  é  $\frac{\underline{c}'B_0\underline{c}}{\underline{c}'\Sigma\underline{c}}$ , que é uma generalização multigrupal do caso de duas populações. Medindo a variabilidade entre grupos de valores (escores)  $Y$  relativamente à variabilidade comum dentro dos grupos, da mesma forma do que no caso de duas populações, pode-se selecionar  $\underline{c}$  que maximiza esta razão. É conveniente normalizar  $\underline{c}$  tal que  $\underline{c}'\Sigma\underline{c} = 1$ .

Sejam  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s \geq 0$  os  $s \leq \min(g-1, p)$  autovalores não-nulos de  $\Sigma^{-1}B_0$  e  $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_s$  os correspondentes autovetores (escalonados tais que  $\underline{c}'\Sigma\underline{c} = 1$ ). Então, é fácil provar que o vetor de coeficientes  $\underline{c}$  que maximiza  $\frac{\underline{c}'B_0\underline{c}}{\underline{c}'\Sigma\underline{c}}$  é dado por  $\underline{c}_1 = \underline{e}_1$ .

A combinação linear  $\underline{c}_1' \underline{X}$  é chamada de primeira discriminante e de forma idêntica, pode-se generalizar para o  $k$ -ésimo discriminante com  $\underline{c}_k = \underline{e}_k$ , com  $k = 1, 2, \dots, s$ . Como geralmente  $\Sigma$  e  $\underline{\mu}_i$  não são conhecidas, tomam-se amostras aleatórias de tamanho  $n_i$  das populações  $\Pi_i, i = 1, 2, \dots, g$  e denotamos o conjunto de dados da população  $\Pi_i, i = 1, 2, \dots, g$ , por  ${}_{n_i}Y_p$  tem-se os estimadores dos parâmetros  $\underline{\mu}_i$  e  $\bar{\underline{\mu}}$  dados por:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij} \quad (2.43)$$

$$\bar{\underline{x}} = \frac{\sum_{i=1}^g n_i \bar{x}_i}{\sum_{i=1}^g n_i} = \frac{\sum_{i=1}^g \sum_{j=1}^{n_i} x_{ij}}{\sum_{i=1}^g n_i} \quad (2.44)$$

A Matriz “Soma de produtos cruzados entre grupos”  $B_0$  é estimada por:

$$\hat{B}_0 = \sum_{i=1}^g (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \quad (2.45)$$

e um estimador para  $\Sigma$  pode ser obtido com base na matriz  $W$ .

$$W = \sum_{i=1}^g \sum_{j=1}^{n_i} (x_{ij} - \bar{x})(x_{ij} - \bar{x})' = \sum_{i=1}^g (n_i - 1)S_i \quad (2.46)$$

Consequentemente,

$$\frac{W}{n_1 + n_2 + \dots + n_g - g} = \frac{(n_1 - 1)S_1 + (n_2 - 1)S_2 + \dots + (n_g - 1)S_g}{n_1 + n_2 + \dots + n_g - g} = S_p \quad (2.47)$$

Assim, o mesmo  $\hat{c}$  que maximiza a razão  $\frac{\hat{c}' B_0 \hat{c}}{\hat{c}' S_p \hat{c}}$  também maximiza  $\frac{\hat{c}' B_0 \hat{c}}{\hat{c}' W \hat{c}}$ .

Logo, apresentar-se-á otimizante  $\hat{c}$  na forma mais usual, que é o autovalor  $\hat{e}_i$  da matriz  $W^{-1}B_0$ , porque se  $W^{-1}B_0\hat{e} = \hat{\lambda}\hat{e}$  então  $S_p^{-1}\hat{B}_0\hat{e} = \hat{\lambda}(n_1 + n_2 + \dots + n_g - g)\hat{e}$ , portanto, concluindo que sejam  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_s > 0$  os autovalores não nulos de  $W^{-1}B_0$  e  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_s$  os correspondentes autovalores, sendo  $s \leq \min(g-1, p)$  e  $\hat{e}_i$  normalizado tal que  $\hat{e}_i' S_p \hat{e}_i = 1$ ; então o vetor de coeficientes que maximiza a razão citada acima é  $\hat{c}_1 = \hat{e}_1$  e a combinação linear  $\hat{e}_1' x$  é chamada primeira discriminante amostral. Generalizando, tem-se no passo  $k$  o  $k$ -ésimo discriminante amostral  $\hat{e}_k' x$ ,  $k \leq s$ .

### 2.2.1.5 Escore Discriminante Quadrático de Classificação

Quando a hipótese nula,  $H_0$ , enunciada no item anterior não é aceita, uma das alternativas é optar-se pela técnica do escore discriminante quadrático de classificação. Considere o vetor de variáveis aleatórias  $X$  de populações com distribuição normal multivariada com média  $\mu_i$  e matriz de covariância  $\Sigma_i$ , ou seja,

$$f_i(\underline{X}) = \frac{1}{(2\pi)^{p/2} |\Sigma_i|} \exp \left[ -\frac{1}{2} (\underline{X} - \underline{\mu}_i)' \Sigma_i (\underline{X} - \underline{\mu}_i) \right], \quad i = 1, 2, \dots, g \quad (2.48)$$

em que  $\underline{\mu}_i = E(\underline{x} | \Pi_i)$  é o vetor de médias e  $V(\underline{x} | \Pi_i) = \Sigma_i$  é a matriz de covariância.

Tem-se ainda que os custos  $c(i|i) = 0$  e  $c(k|i) = 1, k \neq i, k = 1, 2, \dots, g$  e da equação (2.7), a regra de classificar fica:

alocar  $\underline{X}$  em  $\Pi_k$  se

$$\ln p_i f_i(\underline{X}) = \ln p_i - \left( \frac{p}{2} \right) \ln(2\pi) - \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (\underline{X} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{X} - \underline{\mu}_i), \quad i = 1, 2, \dots, g \quad (2.49)$$

Na equação anterior pode-se ignorar a constante  $\left( \frac{p}{2} \right) \ln(2\pi)$  já que é a mesma para todas as populações. Define-se, então, o escore discriminante quadrático de classificação  $d_i^Q$  para a população  $\Pi_i$  como:

$$d_i^Q(\underline{X}) = \frac{1}{2} \ln |\Sigma_i| - \frac{1}{2} (\underline{X} - \underline{\mu}_i)' \Sigma_i^{-1} (\underline{X} - \underline{\mu}_i) + \ln p_i, \quad i = 1, 2, \dots, g \quad (2.50)$$

Com várias populações normais, a regra de classificação consiste em classificar  $\underline{X}$  em  $\Pi_k$  se:

$$d_i^Q(\underline{X}) = \max_i d_i^Q(\underline{X}), \quad i = 1, 2, \dots, g \quad (2.51)$$

Considerando que os parâmetros populacionais em geral não são conhecidos, as equações com os estimadores serão:

$$D_i = \frac{1}{2} \ln |S_i| - \frac{1}{2} (\underline{X} - \bar{\underline{X}}_i)' S_i^{-1} (\underline{X} - \bar{\underline{X}}_i) + \ln p_i, \quad i = 1, 2, \dots, g \quad (2.52)$$

onde:

- $\bar{\underline{X}}_i$  é o vetor médio do grupo  $i$ ,
- $S_i$  é a matriz de covariância do grupo  $i$ .
- $P_i$  é a probabilidade a priori de um individuo ser classificado no grupo  $i$ .

### 2.2.2 Regressão Logística

A função Logística surgiu em 1845, ligada a problemas de crescimento demográfico. Em tais aspectos esta função desempenha um papel importante. A partir da década de 30 esta metodologia passou a ser aplicada no âmbito da biologia Lima (2002).

Este método, assim como a regressão linear e múltipla, estuda a relação entre uma variável resposta e uma ou mais variáveis independentes. A diferença entre estas técnicas de regressão se deve ao fato de que na regressão logística as variáveis dependentes estão dispostas em categorias discriminatórias enquanto na regressão linear estas variáveis são dados contínuos. Outra diferença é que na regressão logística a resposta pode ser entendida como uma probabilidade de ocorrência  $p \in [0,1]$ , enquanto na regressão linear obtém-se um valor numérico  $p \in \mathcal{R}$ .

#### 2.2.2.1 Modelo de Regressão Logística

Seja uma variável aleatória  $Y$  (dicotômica) e  $\underline{X}' = [X_1, X_2, \dots, X_p]$  um vetor de dimensão  $p$ , composto de variáveis aleatórias independentes e ainda tomando-se  $n$  observações independentes de  $Y$  e  $X_i$  com  $i=1, 2, \dots, p$ , pode-se escrever o modelo de Regressão Logística, na forma:

$$P(\tilde{X}) = \frac{e^{\underline{\beta}'\underline{X}}}{1 + e^{\underline{\beta}'\underline{X}}} = \frac{1}{1 + e^{-\underline{\beta}'\underline{X}}} \quad (2.53)$$

onde:  $\underline{\beta}' = [\beta_0, \beta_1, \dots, \beta_p]$  e  $\underline{X}' = [1, X_1, \dots, X_p]$

Segundo Johson & Wichern, (1998), esse modelo é o mais apropriado no caso da variável dependente ser dicotômica, pois a Regressão Logística estima diretamente a probabilidade de ocorrência de um evento. Estes autores, também salientam que a estimação dos parâmetros deve ser feita pelo método da máxima verossimilhança, sendo o mais recomendado quando dispõe-se de observações individuais da ocorrência ou não de determinado evento.



### 2.2.2.2 Modelo de Regressão Logística Simples

Seja a amostra aleatória composta de  $n$  pares de observações  $(X_i, Y_i)$  com  $i = 1, 2, \dots, n$ , onde os  $Y$ 's representam os valores observados de uma variável dicotômica, e os  $X$ 's representam os valores observados de uma única variável independente.

Assim o modelo  $P(X) = \frac{e^{\frac{\beta X}{1 + e^{\frac{\beta X}{1 + e^{\beta_0 + \beta_1 X}}}}}}{1 + e^{\frac{\beta X}{1 + e^{\beta_0 + \beta_1 X}}}}$ , torna-se:

$$P(Y = 1 | X) = P(X) = \frac{e^{\beta_0 + \beta_1 X}}{1 + e^{\beta_0 + \beta_1 X}} = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X)}} \quad (2.54)$$

Essa expressão fornece a probabilidade condicional de que  $Y$  seja igual a 1, dado o valor de  $X$ , ou seja,  $P(Y = 1 | X)$ . Neste caso é necessário estimar os valores dos parâmetros  $\beta$ 's. Então com isso determinando o modelo de Regressão Logística a probabilidade condicional de que  $Y$  ser igual a 0, é:

$$P(Y = 0 | X) = 1 - P(Y = 1 | X) \quad (2.55)$$

Portanto para as  $n$  observações têm-se:

$$P(Y_i = 1 | X_i) = P(X_i) \quad e \quad P(Y_i = 0 | X_i) = 1 - P(X_i) \quad (2.56)$$

O método da Máxima Verossimilhança é melhor método para se estimar os parâmetros do modelo de Regressão Logística, ou seja, fornece estimativas para os parâmetros que maximizam a probabilidade de obter o conjunto observado de dados.

Uma forma conveniente para expressar a contribuição da função de verossimilhança para os pares  $(X_i, Y_i)$  é através da fórmula:

$$F(X_i) = P(X_i)^{Y_i} [1 - P(X_i)]^{1 - Y_i} \quad (2.57)$$

Desde que assumindo que as observações são independentes, a função de verossimilhança é obtida como o produto dos termos dados na equação (2.54) como segue:

$$l(\beta) = \prod_{i=1}^n F(X_i) = \prod_{i=1}^n [P(X_i)^{Y_i}][1 - P(X_i)]^{1-Y_i} \quad (2.58)$$

Os estimadores de Máxima Verossimilhança destes parâmetros são escolhidos de forma a maximizar essa função, ou seja, deseja-se determinar o estimador para  $\beta$  que maximize a expressão anterior. Analisando essa função do ponto de vista matemático é mais fácil trabalhar com o logaritmo da mesma. Assim, tem-se a função de log-verossimilhança dada por:

$$L(\beta) = \ln(l(\beta)) = \ln \left[ \prod_{i=1}^n F(X_i) \right] = \ln \left[ \prod_{i=1}^n [P(X_i)^{Y_i}][1 - P(X_i)]^{1-Y_i} \right] \quad (2.59)$$

Assim:

$$L(\beta) = \sum_{i=1}^n [Y_i \ln(P(X_i))] + (1 - Y_i) \ln[1 - P(X_i)] \quad (2.60)$$

Então:

$$L(\beta) = \sum_{i=1}^n \left[ Y_i \ln \frac{e^{\beta'X}}{1 + e^{\beta'X}} + (1 - Y_i) \ln \left( 1 - \frac{e^{\beta'X}}{1 + e^{\beta'X}} \right) \right] \quad (2.61)$$

Temos:

$$L(\beta) = \sum_{i=1}^n \left[ Y_i \ln \frac{e^{\beta'X}}{1 + e^{\beta'X}} + (1 - Y_i) \ln \left( \frac{1}{1 + e^{\beta'X}} \right) \right] \quad (2.62)$$

Mas

$$\ln \left( \frac{e^{\beta'X}}{1 + e^{\beta'X}} \right) = \ln e^{\beta'X} - \ln(1 + e^{\beta'X}) = \beta'X - \ln(1 + e^{\beta'X}) \quad (2.63)$$

e

$$\ln \left( \frac{1}{1 + e^{\beta'X}} \right) = \ln 1 - \ln(1 + e^{\beta'X}) = -\ln(1 + e^{\beta'X}) \quad (2.64)$$

Substituindo (2.63) e (2.64) em (2.62) temos:

$$L(\beta) = \sum_{i=1}^n [Y_i(\underline{\beta}' X - \ln(1 + e^{\underline{\beta}' X})) + (1 - Y_i)(-\ln(1 + e^{\underline{\beta}' X}))] \quad (2.65)$$

Assim:

$$L(\beta) = \sum_{i=1}^n [Y_i \underline{\beta}' X - Y_i \ln(1 + e^{\underline{\beta}' X}) - \ln(1 + e^{\underline{\beta}' X}) + Y_i \ln(1 + e^{\underline{\beta}' X})] \quad (2.66)$$

Concluindo:

$$L(\beta) = \sum_{i=1}^n [Y_i \underline{\beta}' X - \ln(1 + e^{\underline{\beta}' X})] \quad (2.67)$$

É possível demonstrar a concavidade global da função de log-verossimilhança sendo, portanto, apenas necessário anular o vetor das primeiras derivadas desta função em relação a  $\beta$  para obter os estimadores de máxima verossimilhança. Assim para determinar o valor de  $\beta$  que maximiza  $L(\beta)$  deve-se derivar  $L(\beta)$  em relação à  $\beta_0$  e  $\beta_1$  e igualar o conjunto de resultados a zero. Desta forma tem-se a seguir as chamadas equações de verossimilhança, que são derivadas da expressão anterior em relação à  $\beta$ .

$$\frac{\partial L(\beta)}{\partial \beta_0} = \sum_{i=1}^n \left( Y_i - \frac{e^{\underline{\beta}' X}}{1 + e^{\underline{\beta}' X}} \right) = \sum_{i=1}^n [Y_i - P(X_i)] = 0 \quad (2.68)$$

$$\frac{\partial L(\beta)}{\partial \beta_1} = \sum_{i=1}^n \left( X_i \cdot Y_i - X_i \frac{e^{\underline{\beta}' X}}{1 + e^{\underline{\beta}' X}} \right) = \sum_{i=1}^n X_i [Y_i - P(X_i)] = 0 \quad (2.69)$$

Temos que:

$$\sum_{i=1}^n Y_i = \sum_{i=1}^n \frac{e^{\underline{\beta}' X}}{1 + e^{\underline{\beta}' X}} = \sum_{i=1}^n P(X_i) \quad (2.70)$$

O que reflete no fato da soma dos valores observados de  $Y$  ser igual a soma dos valores esperados (preditos), mediante a aplicação do modelo obtido. Esta propriedade é importante, pois ajuda avaliar o ajuste do modelo.

O valor de  $\beta$  dado pela solução das equações anteriores é chamado de estimador de máxima verossimilhança, é denotado por  $\hat{\beta} \cdot P(X_i)$ , e representa uma estimativa da probabilidade condicional que  $Y$  é igual a 1, são que  $X$  é igual a  $X_i$ .

### 2.2.2.3 Modelo de Regressão Logístico Múltiplo

Considerando-se uma amostra de  $n$  pares de observações independentes  $(\underline{X}_i, Y_i)$ ,  $i = 1, 2, \dots, p$ , da mesma forma como ocorreu no caso univariado, item anterior, o ajuste do modelo requer que se obtenha uma estimativa do vetor  $\underline{\beta}' = [\beta_0, \beta_1, \dots, \beta_p]$ . O método de estimação usado no caso multivariado será o mesmo no univariado, ou seja, o da máxima verossimilhança. A função de máxima verossimilhança é aproximadamente idêntica a dada na equação com uma mudança, sendo que  $P(X)$  é definido por:

$$P(Y = 1 | \underline{X}) = P(\underline{X}_i) = \frac{e^{\underline{\beta}' \underline{X}}}{1 + e^{\underline{\beta}' \underline{X}}} \quad (2.71)$$

Estimar-se-á  $(p+1)$  equações de verossimilhança as quais são obtidas por derivadas da função log-verossimilhança com respeito à  $(p+1)$  parâmetros. As equações de verossimilhança resultantes podem ser representadas como:

$$\sum_{i=1}^n [Y_i - P(X_i)] = 0 \quad (2.72)$$

e

$$\sum_{i=1}^n X_{ij} [Y_i - P(X_i)] = 0 \quad (2.73)$$

para  $j = 1, 2, \dots, p$

Da mesma forma como ocorreu com o modelo univariado, as soluções das equações de verossimilhança requer a utilização de um método iterativo na busca dos parâmetros estimados  $\hat{\beta}'$ .

## 2.2.3 Redes Neurais Artificiais (RNAs)

### 2.2.3.1 Definição de RNAs

Muitas definições foram elaboradas desde que as RNAs começaram a surgir como uma nova perspectiva à Inteligência Artificial. Em todas elas prevaleceram às semelhanças ao sistema nervoso humano. A maioria das pesquisas sobre RNAs foram motivadas pela idéia de que o cérebro humano é um computador com processamento de informações bastante complexas.

Haykin (2002) define as redes neurais como sendo: Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem a propensão natural para armazenar conhecimento experimental e torná-lo disponível para uso. Ela se assemelha ao cérebro em dois aspectos:

1. O de que o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizagem.
2. Forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido.

Dessa forma, uma RNA pode ser definida como uma estrutura computacional que tem como objetivo permitir a implementação de modelos matemáticos que representem a forma como o cérebro humano processa as informações que adquire.

Apesar de muitas definições darem a impressão de simplicidade, é de fundamental importância ressaltar que na prática, essa idéia não acontece de forma tão simples. Durante a implementação da RNA vários fatores podem ocorrer impossibilitando a aplicação em certos casos. Percebe-se também que as soluções não são precisas quando as RNAs são aplicadas isoladamente.

Também não se pode deixar de mencionar que ainda existe uma longa e árdua trajetória a ser percorrida no sentido das pesquisas para chegar a uma arquitetura computacional que reproduza fielmente a capacidade do cérebro humano.

### 2.2.3.2 Neurônio Artificial

O neurônio artificial é a estrutura básica de processamento de uma RNA. Os neurônios artificiais são estruturas lógicas matemáticas que têm o objetivo de simular o funcionamento de um neurônio biológico, fazendo com que a RNA atue de forma indutiva. A figura 2.8 mostra um modelo de neurônio artificial, com suas entradas ( $X_1, X_2, \dots, X_n$ ), pesos sinápticos ( $W_{1k}, W_{2k}, \dots, W_{nk}$ ), bias ( $b_k$ ), função soma ( $\Sigma$ ), função de transferência ( $\varphi$ ) e saída ( $a_k$ ).

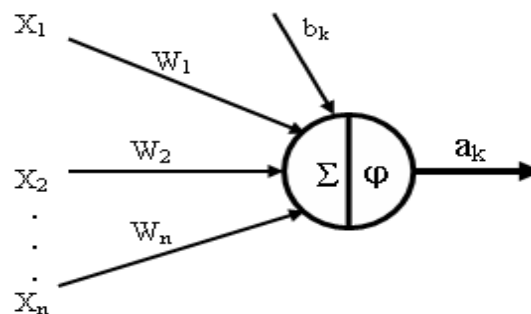


FIGURA 2.2 – MODELO DE UM NEURÔNIO ARTIFICIAL  
FONTE: O AUTOR

O neurônio funciona basicamente como integrador de sinais. Tem a função de coletar sinais provenientes de outros neurônios, ou sinais da entrada da rede, ponderar os dados através de pesos específicos e repassar esses dados para outros neurônios ou para saída da rede. De acordo com a figura 2. pode-se entender a estrutura básica de um neurônio artificial.

Para se calcular o sinal de saída ( $a_k$ ), efetuada por um neurônio desse tipo, pode-se utilizar a equação a seguir:

$$a_k = \varphi \left( \sum_{i=1}^n x_i w_{ik} + b_k \right) \quad (2.74)$$

onde :

- $n$  = o número total de sinais de entrada que chegam até o neurônio que está sendo calculado;
- $x_i$  = valor de entrada para esse neurônio;
- $w_{ik}$  = peso associado a essa conexão;
- $b_k$  = têm a função de aumentar ou diminuir a entrada da função de ativação.

O resultado da somatória das componentes do produto interno entre o vetor de entrada ( $X$ ) e o vetor de pesos ( $W$ ) do neurônio, é mapeado através da função de ativação ( $\varphi$ ), produzindo o valor da saída ( $a_k$ ) do neurônio ( $k$ ).

### 2.2.3.3 Funções de Ativação de uma RNA

O objetivo de uma função de ativação é restringir a amplitude da saída do neurônio, ou seja, limitar sua saída em um intervalo de valor finito.

A função de ativação que define a saída do neurônio artificial, ou seja, o caminho onde a informação será conduzida, tornando-se muito importante para o comportamento da RNA.

Os tipos de função de ativação mais usados para mapear uma RNA são basicamente três: Limiar ou Degrau, Linear por partes e Sigmoidal. As representações gráficas dessas funções de ativação são mostradas na figura 2.9, a seguir.

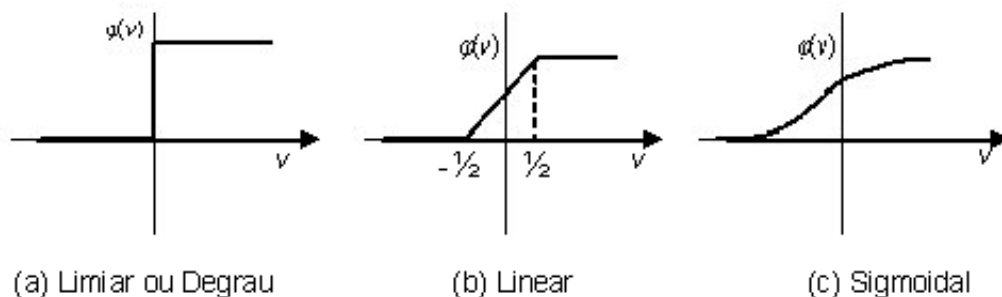


FIGURA 2.3 – FUNÇÕES DE ATIVAÇÃO DE UMA RNA  
 FONTE: O AUTOR

### 2.2.3.4 Características Gerais das RNAs

Uma das principais vantagens e importância das RNAs é a sua capacidade de “aprender sozinha” através do seu próprio uso, e de melhorar seu desempenho baseado nessa aprendizagem.

A capacidade de uma rede aprender vem dos neurônios artificiais, que têm funcionamento baseado no neurônio biológico. Embora pareça simples, o

aprendizado é um processo altamente complexo. Além de envolver diversos níveis de abstração, também reúne habilidades e conceitos advindos de várias áreas do conhecimento e não há em relação a ele uma definição única e plena.

Apesar da complexidade, a analogia com os elementos do cérebro humano é essencial para entender-se de maneira satisfatória a estrutura e o funcionamento de uma RNA.

Segundo Pandya e Macy (1996), o sistema nervoso de um ser humano, assim como o de qualquer primata, consiste de três estágios básicos: o estágio inicial, no qual os estímulos de entrada externos são captados pelas unidades "sensoras" (os olhos, por exemplo, captam sinais visuais); o estágio de processamento destes sinais pelo sistema nervoso central (o cérebro) e o estágio no qual as respostas aos estímulos de entrada são obtidas (a voz pode ser ativada em resposta a determinada entrada, bem como ações motoras ou orgânicas).

De acordo com esta definição, o cérebro humano recebe informações do meio exterior, processa, avalia e compara as outras informações que tem armazenado na memória, tornando-se apto a tomar decisões apropriadas. As RNAs são baseadas no funcionamento no cérebro humano. Assim, ela é estruturada de maneira a simular todas as etapas de processamento das informações. Segundo a definição de Pandya e Macy (1996), uma RNA pode apresentar-se da seguinte forma:

- uma camada inicial de unidades sensoras aos estímulos de entrada, que não é composta por neurônios;
- uma ou mais camadas intermediárias de neurônios interligados (formando a estrutura central de processamento da rede neural);
- uma camada de neurônios de saída, conectada aos neurônios da(s) camada(s) escondida(s), gerando a resposta final da RNA.

A figura 2.4, a seguir, mostra a estrutura de uma RNA. As entradas da rede representam a informação existente. Elas alimentam a primeira camada de neurônios que através do cálculo executado pela aplicação da função de transferência, geram repostas que alimentam a camada seguinte, e assim atingem a camada de saída.



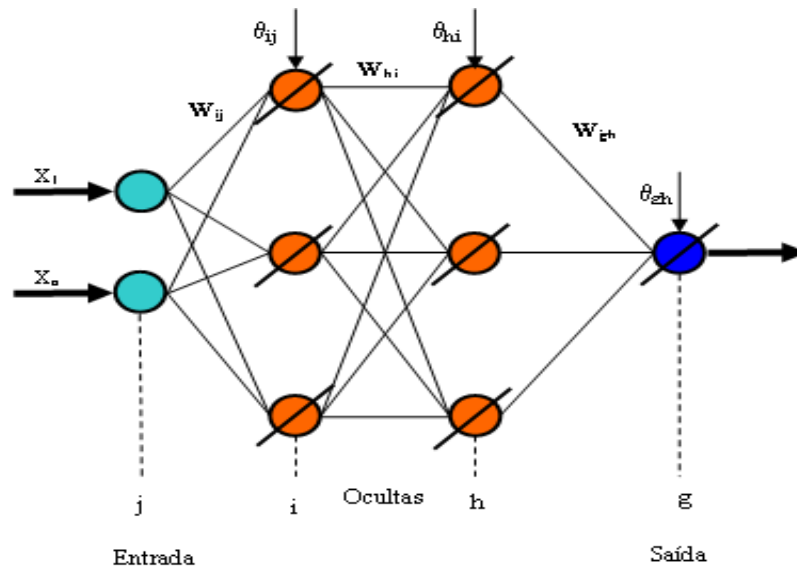


FIGURA 2.4 – REDE NEURAL ARTIFICIAL  
 FONTE: O AUTOR

A forma com que os neurônios se agrupam em camadas e o conjunto de suas conexões define a arquitetura da RNA. As arquiteturas neurais são organizadas em camadas, com os neurônios de uma determinada camada, conectados aos neurônios da camada posterior. Os dispositivos processadores são agrupados por pelo menos duas camadas: uma camada para apresentação dos dados e outra para obtenção dos resultados. As redes, em geral, possuem uma ou mais camadas intermediárias, também conhecidas como camadas escondidas. Segundo Haykin (2002) existem três tipos de arquitetura de rede:

- Redes alimentadas adiante com camada única (*feedforward*): o termo camada única se refere a camada de saída, ou seja, o processamento ocorre somente na camada de saída. A figura a seguir mostra um modelo de uma rede *feedforward* com uma única camada.

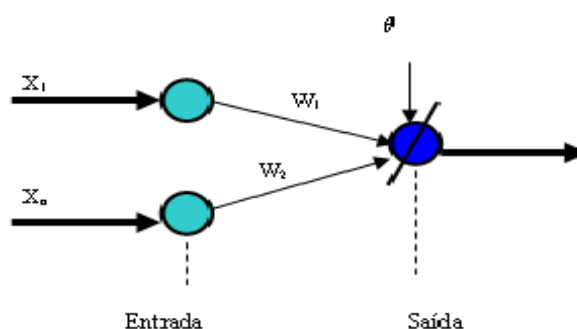


FIGURA 2.5 – MODELO DE REDE NEURAL COM UMA CAMADA  
 FONTE: O AUTOR

- redes alimentadas adiante com múltiplas camadas (*feedforward*): esta possui uma ou mais camadas ocultas. A figura a seguir ilustra uma rede *feedforward* com duas camadas – uma oculta e uma de saída.

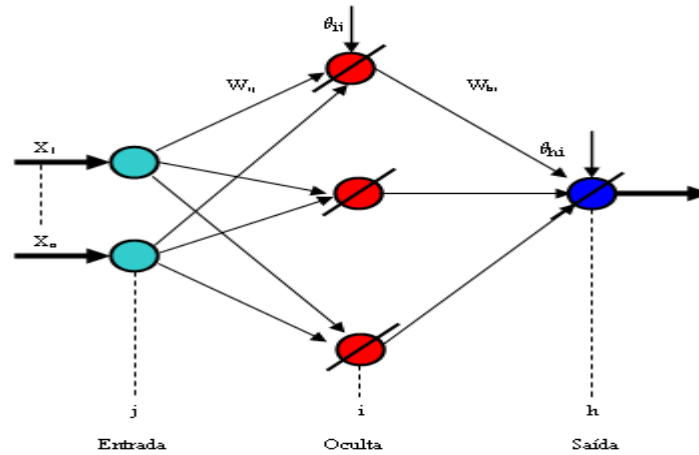


FIGURA 2.6 – MODELO DE REDE NEURAL COM DUAS CAMADAS  
 FONTE: O AUTOR

- redes recorrentes: possui pelo menos um laço de realimentação. A figura a seguir ilustra esse tipo de arquitetura.

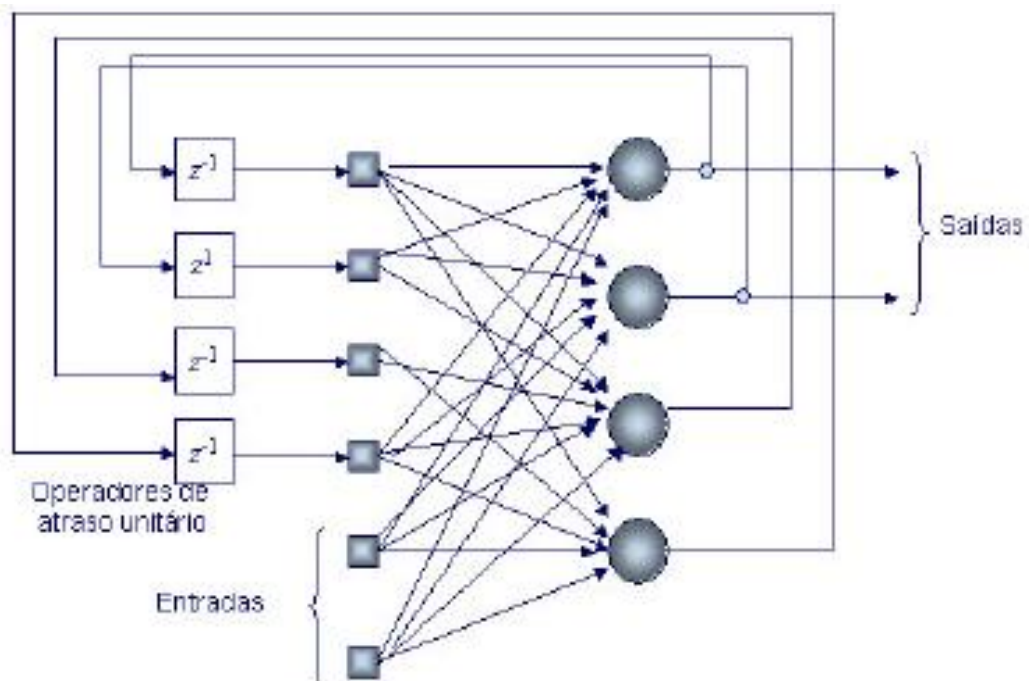


FIGURA 2.7 – MODELO DE REDE NEURAL RECORRENTE  
 FONTE: O AUTOR

Geralmente as redes *feedforward* de múltiplas camadas apresentam a seguinte estrutura:

- Camada de entrada: onde os padrões são apresentados à rede; nesta camada não ocorre o processamento de informações;
- Camadas intermediárias ou escondidas (não obrigatórias): recebem informação da camada anterior, processam os dados e enviam a resposta para a camada seguinte;
- Camada de saída: local onde se obtém o resultado final através do processamento das informações vindas da camada anterior.

#### 2.2.3.5 Treinamento de uma RNA de Múltiplas Camadas

Em geral o treinamento é um processo gradual e lento, onde os pesos são modificados várias vezes, seguindo-se um algoritmo de treinamento que estabelece a forma como estes pesos são alterados. O treinamento de uma RNA pode ser efetuado de uma das seguintes maneiras:

Treinamento supervisionado: é utilizado quando são conhecidos os valores das saídas para um determinado grupo de padrões. Para cada padrão apresentado, a RNA calcula a saída (valor obtido) e o compara com o valor desejado (real) e atualiza os pesos das conexões entre todos os neurônios buscando reduzir o erro. Redes de Retropropagação, *GRNN* (“*General Regression Neural Networks*”), *PNN* (“*Probabilistic Neural Networks*”), e *GMDH* (“*Group Method Data Handling*”) são exemplos de redes supervisionadas.

Treinamento não supervisionado: neste caso, como não se conhecem as saídas, os pesos são modificados em função de critérios internos.

#### 2.2.3.6 Algoritmo *Back- Propagation*

Segundo Haykin (2002) e Fausett (1994), o algoritmo *back-propagation* é o algoritmo de aprendizado supervisionado mais conhecido. Este algoritmo é de fácil compreensão, o que favorece a sua utilização.

Durante o treinamento com este algoritmo, a rede opera em dois estágios fundamentais. No primeiro, *forward* (para frente), deve-se apresentar um padrão à camada de entrada. Todo o processo é realizado, camada por camada, até que se obtenha uma resposta na camada de saída. Neste momento a saída obtida é comparada à saída desejada e calcula-se o erro. No estágio seguinte, *backward* (para trás), a propagação ocorre da camada de saída para a camada de entrada, onde os pesos das conexões das camadas internas vão sendo modificados conforme o erro é retro propagado.

O ajuste corretivo nos pesos das conexões entre os neurônios, para se obter uma boa aproximação da saída encontrada à resposta desejada, será realizado passo a passo, até que se atinja uma situação estável, ou seja, quando os pesos não precisam mais ser ajustados.

## 2.3 AVALIAÇÃO DA FUNÇÃO DE CLASSIFICAÇÃO

### 2.3.1 Critério TPM (*Total Probability of Misclassification*)

Uma das formas de avaliar o desempenho de um procedimento de reconhecimento de padrões consiste no cálculo da taxa de erro de classificação (*Total Probability of Misclassifications* - TPM) dado por:

$$TPM = p_1 \int_{R_2} f_1(\underline{X})d\underline{X} + p_2 \int_{R_1} f_2(\underline{X})d\underline{X} \quad (2.75)$$

onde  $p_1$  e  $p_2$  são as probabilidades de uma observação pertencente a  $\Pi_1$  ou a  $\Pi_2$ , respectivamente.

O valor mínimo dessa quantidade, obtido pela escolha adequada das regiões  $R_1$  e  $R_2$ , é chamado de taxa ótima de erro (*Optimum Error Rate* - OER). A OER é dada por:

$$OER = p_1 \int_{R_2} f_1(\underline{X})d\underline{X} + p_2 \int_{R_1} f_2(\underline{X})d\underline{X} \quad (2.76)$$

onde as regiões são determinadas por:

$$R_1 = \frac{f_1(\underline{X})}{f_2(\underline{X})} \geq \frac{p_2}{p_1} \text{ e } R_2 = \frac{f_1(\underline{X})}{f_2(\underline{X})} < \frac{p_2}{p_1} \quad (2.77)$$

Assim a OER é a taxa de erro mínima para a regra de classificação TPM. O desempenho da função de classificação da amostra pode, a princípio, ser avaliada pelo cálculo da taxa real de erro (*Actual Error Rate - AER*)

$$AER = p_1 \int_{\hat{R}_2} f_1(\underline{X}) d\underline{X} + p_2 \int_{\hat{R}_1} f_2(\underline{X}) d\underline{X} \quad (2.78)$$

onde:  $\hat{R}_1$  e  $\hat{R}_2$  representam as regiões de classificação determinadas pelos tamanhos das amostras  $n_1$  e  $n_2$ , respectivamente.

A AER indica o desempenho da função de classificação da amostra em amostras futuras. Da mesma forma que a taxa OER, não pode, em geral, ser calculada porque depende das f.d.p's  $f_1(\underline{X})$  e  $f_2(\underline{X})$  que são desconhecidas. Entretanto uma estimativa da quantidade ligada à taxa de erro real pode ser calculada, e sua estimativa será discutida na seqüência.

Há medidas de desempenho que não dependem das f.d.p's e podem ser calculadas por qualquer procedimento de classificação. Por exemplo, a medida, chamada de taxa de erro aparente (*Apparent Error Rate - APER*), é definida como a fração das observações usadas no treinamento das amostras que são incorretamente classificadas pela função de classificação amostral.

A APER pode ser facilmente calculada pela matriz de confusão, a qual mostra a classificação real versus a classificação prevista. Para  $n_1$  observações de  $\Pi_1$  e  $n_2$  observações de  $\Pi_2$ , a matriz de confusão tem a forma dada na tabela 2.1.

TABELA 2.1 – TABELA DE CLASSIFICAÇÃO

Classificação Prevista		
Classificação	$\Pi_1$	$\Pi_2$
Real		
$\Pi_1$	$n_{1/1}$	$n_{1/2}$
$\Pi_2$	$n_{2/1}$	$n_{2/2}$

FUNTE: APPLIED MULTIVARIATE STATISTICAL ANALYSIS

onde:

- $n_{i/j}$  = números de observações de  $\Pi_i$  classificadas em  $\Pi_j$  e consideradas corretas se  $i = j$  ou incorreta se  $i \neq j$ ;
- $n_{1/1}$  = numero de itens de  $\Pi_1$  classificadas corretamente como item de  $\Pi_1$ ;
- $n_{1/2}$  = número de itens de  $\Pi_1$  classificados incorretamente como item de  $\Pi_2$ ;
- $n_{2/2}$  = número de itens de  $\Pi_2$  classificados corretamente como item de  $\Pi_2$ ;
- $n_{2/1}$  = número de itens de  $\Pi_2$  classificados incorretamente como item de  $\Pi_1$ .

A taxa aparente de erro (*Apparent Error Date* – APER) é dada por:

$$APER = \frac{n_{1/2} + n_{2/1}}{n_1 + n_2} \quad (2.79)$$

Essa taxa deve ser interpretada como a proporção de observações incorretamente classificadas.

A APER é intuitivamente atraente e de cálculo fácil. Infelizmente, é tendenciosa para subestimar a AER, e o problema não desaparece a menos que os tamanhos das amostras  $n_1$  e  $n_2$  sejam grandes. Essencialmente, uma avaliação ótima ocorre porque os dados usados para construir a função de classificação são também usados para sua avaliação.

A taxa de erro estimada pode ser contida, pois são melhores que a taxa de erro aparente, mantendo-se relativa facilidade de cálculo, e não requer a função de distribuição de probabilidade. Um procedimento é dividir a amostra total em uma parte da amostra para treinamento e outra parte para validação. A parte da amostra

para um treinamento é usada para construir a função de classificação, e a parte para validação é usada para sua avaliação.

A taxa de erro é determinada pela proporção de classificação errada dentro da amostra considerada. Embora este método supere o problema da tendenciosidade por não usar os mesmos dados para construir e julgar a função de classificação, ele sofre de dois defeitos principais:

É necessário uma grande amostra;

A função de avaliação não é a função de interesse. E ainda, quase todos os dados devem ser usados para a construção da função de classificação, pois o valor da informação pode ser perdido.

Uma segunda abordagem que funciona muito bem é o procedimento proposto por Lachenbruch (1975).

### 2.3.2 Abordagem de Lachenbruch

O método de Lachenbruch, segundo Johnson & Wichern (1998), é uma abordagem estatística empregada para avaliar o desempenho da regra de reconhecimento de padrão e estimar a probabilidade de classificação correta e incorreta em cada grupo, ou seja, uma forma de avaliar a eficiência da regra de classificação.

O algoritmo do referido procedimento é:

1. Inicie com as observações do grupo  $\Pi_1$ , omita uma observação desse grupo, e construa uma função de classificação usando as  $(n_1 - 1)$  e  $n_2$  observações.
2. Classifique a observação retirada, usando a função construída no passo 1.
3. Repita o passo 1 e 2 até que todas as observações de  $\Pi_1$  sejam classificadas.

Seja  $n_{1/2}$  o número de observações reconhecidas erroneamente nesse grupo. Repita os passos de 1 a 3 para as  $n_2$  observações do grupo  $\Pi_2$ .

Seja  $n_{2/1}$  o número de observações reconhecidas erroneamente nesse grupo. Assim as estimativas das probabilidades de classificação incorreta são:

$$P(2/1) = \int_{R_2} f_1(\underline{X}) d\underline{X} = \frac{n_{2/1}}{n_1} \quad (2.80)$$

classificar em  $\Pi_2$ , quando a observação pertence a  $\Pi_1$ ,

$$P(1/2) = \int_{R_1} f_2(\underline{X}) d\underline{X} = \frac{n_{1/2}}{n_2} \quad (2.81)$$

classificar em  $\Pi_1$ , quando a observação pertence a  $\Pi_2$ ,

Desta forma a proporção esperada de erro (probabilidade estimada) é dada por:

$$\hat{E}(AER) = \frac{n_{1/2} + n_{2/1}}{n_1 + n_2} \quad (2.82)$$

## 2.4 ANÁLISE FATORIAL

A Análise Fatorial foi introduzida por Spearman (1904). Foi inicialmente usada nas áreas de psicologia e ciências sociais, na tentativa de identificar os fatores relacionados com a inteligência humana e ligá-los, de algum modo, à etnia. Devido a subjetividade e falta de unicidade de suas soluções, tem sido alvo de críticas ao longo dos anos (HILLS; KACHIGAN, 1977, 1991 apud MINGOTI, 2005). Porém, é uma das técnicas multivariadas mais conhecidas e tem sido muito utilizada na química (OZERENKO; GYUL'MALIEV; GAGARIN, 2007), na educação (MARQUES, 2010a), na geologia (QUEIROZ et al., 2008), na economia (ESQUARCINI, 2005), na agricultura (CHIGUTI, 2005), na engenharia de produção (MÜLLER; CHAVES NETO, 2007), entre outras.



Segundo Hair et al. (2009), a Análise Fatorial é um método estatístico multivariado cujo propósito principal é definir a estrutura subjacente em uma matriz de dados. Em termos gerais, a Análise Fatorial aborda o problema de analisar a estrutura das interrelações (correlações) entre um grande número de variáveis, definindo um conjunto de dimensões latentes comuns, chamadas de fatores. O objetivo principal da Análise Fatorial é resumir os dados, obtendo dimensões latentes, que quando interpretadas e compreendidas, descrevem os dados em um número muito menor de conceitos do que as variáveis originais individuais.

Em linhas gerais Mingoti (2005) afirma que, o que se espera é que as variáveis originais estejam agrupadas em subconjuntos de novas variáveis mutuamente não correlacionadas. Deste modo, em casos no qual se tem um grande número de variáveis medidas e correlacionadas entre si, seria possível, a partir da Análise Fatorial, identificar um número menor de novas variáveis alternativas, não correlacionadas e que de algum modo sumarizassem as informações principais das variáveis originais. A

Análise Fatorial pode ser trabalhada de forma exploratória ou confirmatória. No primeiro caso, busca-se encontrar os fatores subjacentes às variáveis originais amostradas, sem ter noção de quantos fatores serão obtidos e de que maneira as variáveis estarão agrupadas. Na Análise Fatorial confirmatória, tem-se um modelo inicial pré-estabelecido, onde se deseja verificar se este modelo proposto é consistente com os dados amostrais disponíveis.

#### 2.4.1 Modelo Fatorial Ortogonal

De acordo com Johnson e Wichern (2007), o vetor aleatório  $\underline{X}$ , com  $p$  variáveis aleatórias observáveis, tem vetor de médias  $\underline{\mu}$  e matriz de covariância  $\Sigma$ . O modelo fatorial é linearmente dependente sobre variáveis aleatórias não observáveis  $F_1, F_2, \dots, F_m$  ( $m \leq p$ ) chamados fatores comuns e  $p$  fontes de variações ativas  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p$ , chamadas de erros ou fatores específicos. O modelo fatorial é



$$\Sigma = V(\underline{X}) = E[(\underline{X} - \underline{\mu})(\underline{X} - \underline{\mu})'] \quad (2.88)$$

porém de (2.86) obtém-se

$$\begin{aligned} \Sigma &= E[(\underline{L}\underline{F} + \underline{\varepsilon})(\underline{L}\underline{F} + \underline{\varepsilon})'] \\ \Sigma &= E[\underline{L}\underline{F}\underline{F}'\underline{L}' + \underline{\varepsilon}\underline{F}'\underline{L}' + \underline{L}\underline{F}\underline{\varepsilon}' + \underline{\varepsilon}\underline{\varepsilon}'] \\ \Sigma &= \underline{L}E(\underline{F}\underline{F}')\underline{L}' + E(\underline{\varepsilon}\underline{F}')\underline{L}' + \underline{L}E(\underline{F}\underline{\varepsilon}') + E(\underline{\varepsilon}\underline{\varepsilon}') \end{aligned} \quad (2.89)$$

logo

$$\Sigma = \underline{L}\underline{L}' + \Psi \quad (2.90)$$

Desta forma, a covariância entre as variáveis e os fatores é calculado por:

$$\begin{aligned} \text{Cov}(\underline{X}, \underline{F}) &= E[(\underline{X} - \underline{\mu})(\underline{F} - E(\underline{F}))'] \\ \text{Cov}(\underline{X}, \underline{F}) &= E[(\underline{L}\underline{F} + \underline{\varepsilon})\underline{F}'] \\ \text{Cov}(\underline{X}, \underline{F}) &= E[\underline{L}\underline{F}\underline{F}' + \underline{\varepsilon}\underline{F}'] \\ \text{Cov}(\underline{X}, \underline{F}) &= \underline{L}E(\underline{F}\underline{F}') + E(\underline{\varepsilon}\underline{F}') \end{aligned} \quad (2.91)$$

ou seja,

$$\text{Cov}(\underline{X}, \underline{L}) = \underline{L} \quad (2.92)$$

Sendo assim, por  $\Sigma = \underline{L}\underline{L}' + \Psi$  obtém-se

$$V(X_i) = \ell_{i1}^2 + \dots + \ell_{im}^2 + \psi_i \quad (2.93)$$

$$\text{Cov}(X_i, X_k) = \ell_{i1}\ell_{k1} + \dots + \ell_{im}\ell_{km} \quad (2.94)$$

e de (2.92) resulta que a covariância entre a  $i$ -ésima variável e o  $j$ -ésimo fator é:

$$\text{Cov}(X_i, F_j) = \ell_{ij} \quad (2.95)$$

A parte da variância que a  $i$ -ésima variável contribui para os  $m$  fatores comuns é chamada de comunalidade, que é dada por:

$$h_i^2 = \ell_{i1}^2 + \dots + \ell_{im}^2 \quad (2.96)$$

e  $\psi_i$  são as variâncias específicas.

### 2.4.2 Método das componentes principais via matriz de covariância

Usando a decomposição espectral pode-se fatorar a matriz de covariância  $\Sigma$ . Sejam  $(\lambda_i, \underline{e}_i)$  pares de autovalores autovetores da matriz  $\Sigma$ , com  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ . Então:

$$\Sigma = \lambda_1 \underline{e}_1 \underline{e}_1' + \lambda_2 \underline{e}_2 \underline{e}_2' + \dots + \lambda_p \underline{e}_p \underline{e}_p' \quad (2.97)$$

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 & & & \\ & \sqrt{\lambda_2} \cdot \underline{e}_2 & & \\ & & \dots & \\ & & & \sqrt{\lambda_p} \cdot \underline{e}_p \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 \\ \sqrt{\lambda_2} \cdot \underline{e}_2 \\ \vdots \\ \sqrt{\lambda_p} \cdot \underline{e}_p \end{bmatrix} \quad (2.98)$$

o qual ajusta a estrutura de covariância prescrita para a análise do modelo fatorial, tendo  $m$  fatores ( $m=p$ ) e a variância específica  $\psi_i = 0$  para todo  $i=1,2,\dots,p$ . A matriz de carga tem a  $j$ -ésima coluna dada por  $\sqrt{\lambda_j} \cdot \underline{e}_j$ . Desta maneira pode-se escrever,

$$\underset{(p \times p)}{\Sigma} = \underset{(p \times p)}{\underline{L}} \cdot \underset{(p \times p)}{\underline{L}'} + \underset{(p \times p)}{\underline{0}} = \underline{L} \underline{L}' \quad (2.99)$$

entretanto nesse caso a representação de  $\Sigma$  é exata, isto é, o número de fatores comuns é igual ao número de variáveis. Na prática deve-se optar por modelos que expliquem a estrutura da covariância usando poucos fatores comuns. Para isso, eliminam-se os últimos  $(p-m)$  autovalores, que pouco contribuem para  $\Sigma$ . Negligenciando esta contribuição, obtém a aproximação

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 & & & \\ & \sqrt{\lambda_2} \cdot \underline{e}_2 & & \\ & & \dots & \\ & & & \sqrt{\lambda_m} \cdot \underline{e}_m \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 \\ \sqrt{\lambda_2} \cdot \underline{e}_2 \\ \vdots \\ \sqrt{\lambda_m} \cdot \underline{e}_m \end{bmatrix} = \underset{p \times m \times p}{\underline{L}} \underset{m \times p}{\underline{L}'} \quad (2.100)$$

A representação anterior considera que os fatores específicos  $\underline{\varepsilon}$  definidos no modelo fatorial ortogonal são de menor importância e podem ser ignorados na fatoração de  $\Sigma$ . Se os fatores específicos forem incluídos no modelo, a contribuição

deles pode ser estimada tomando-se os elementos da diagonal de  $\Sigma - LL'$ , onde  $LL'$  é dado em (2.100).

Considerando-se os fatores específicos, encontramos a aproximação

$$\Sigma = LL' + \Psi \quad (2.101)$$

$$\Sigma = \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 & & & \\ & \sqrt{\lambda_2} \cdot \underline{e}_2 & & \\ & & \dots & \\ & & & \sqrt{\lambda_m} \cdot \underline{e}_m \end{bmatrix} + \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 \\ \sqrt{\lambda_2} \cdot \underline{e}_2 \\ \vdots \\ \sqrt{\lambda_m} \cdot \underline{e}_m \end{bmatrix} + \begin{bmatrix} \psi_1 & 0 & \dots & 0 \\ 0 & \psi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi_p \end{bmatrix} \quad (2.102)$$

onde  $\psi_i = \sigma_i^2 - \sum_{j=1}^m \ell_{ij}^2$ , com  $i = 1, 2, \dots, p$

A representação (2.102) quando aplicada a matriz de covariância amostral  $S$  ou a matriz de correlação amostral  $R$  é conhecida como a solução por componentes principais do modelo fatorial.

#### 2.4.3 Método das componentes principais via matriz de correlação amostral

Realizando a decomposição espectral da matriz de correlação  $R$ , obtêm os pares de autovalores autovetores  $(\lambda_i, \underline{e}_i)$ , onde  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ . Sejam  $m < p$  o número de fatores comuns. Então a matriz estimada das cargas fatoriais  $\hat{\ell}_{ij}$  é dada por:

$$\hat{L} = \begin{bmatrix} \sqrt{\lambda_1} \cdot \underline{e}_1 & & & \\ & \sqrt{\lambda_2} \cdot \underline{e}_2 & & \\ & & \dots & \\ & & & \sqrt{\lambda_m} \cdot \underline{e}_m \end{bmatrix} \quad (2.103)$$

As variâncias específicas estimadas são os elementos da matriz diagonal  $\underline{S} = \hat{L}\hat{L}'$ , então

$$\hat{\Psi} = \begin{bmatrix} \hat{\psi}_1 & 0 & \dots & 0 \\ 0 & \hat{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \hat{\psi}_p \end{bmatrix}, \text{ com } \psi_i = s_{ii}^2 - \sum_{j=1}^m \hat{\ell}_{ij}^2 \quad (2.104)$$

e as comunalidades são estimadas por:

$$\hat{h}_i^2 = \hat{\ell}_{i1}^2 + \dots + \hat{\ell}_{im}^2 \quad (2.105)$$

A matriz residual pode ser obtida pela aproximação de S pela solução por componentes principais, sendo calculada por:

$$S - (\hat{L}\hat{L}' + \hat{\Psi}) \quad (2.106)$$

onde os elementos da diagonal principal são zero, e os outros elementos também são pequenos, quando escolhe-se o número de fatores m apropriado. Analiticamente, tem-se a soma dos quadrados das entradas de

$$S - (\hat{L}\hat{L}' + \hat{\Psi}) \leq \hat{\lambda}_{m+1}^2 + \dots + \hat{\lambda}_p^2 \quad (2.107)$$

Consequentemente, um pequeno valor para o somatório dos quadrados dos autovalores pode ser negligenciado, implicando um pequeno valor para a soma dos quadrados dos erros da aproximação.

Geralmente a contribuição dos primeiros fatores comuns para a variância amostral das variáveis pode ser grandes. A contribuição para a variância amostral  $s_{ii}$  do primeiro fator comum é  $\ell_{i1}$ ,  $i=1,2,\dots,p$ . Para o primeiro fator, a soma dos quadrados das cargas fatoriais é:

$$\hat{\ell}_{11}^2 + \hat{\ell}_{21}^2 + \dots + \hat{\ell}_{p1}^2 = \left( \sqrt{\hat{\lambda}_1} \hat{e}_1 \right)' \left( \sqrt{\hat{\lambda}_1} \hat{e}_1 \right) = \hat{\lambda}_1 \quad (2.108)$$

Desde que  $\hat{e}_1$  tenha tamanho unitário. Em geral a proporção da variância total amostral no  $j$ -ésimo fator é dada por:

$$\left\{ \begin{array}{ll} \frac{\hat{\lambda}_j}{s_{11} + s_{22} + \dots + s_{pp}} & \text{para análise fatorial de S} \\ \frac{\hat{\lambda}_j}{p} & \text{para análise fatorial de R} \end{array} \right. \quad (2.109)$$

O critério dado (anterior) geralmente é usado para determinar o número de fatores comuns apropriado. Quanto maior o número de fatores, maior é a proporção da variância amostral explicada pelo modelo. Existem algumas maneiras de determinar o número de fatores adequados, pode-se destacar o critério de Kaiser e o Scree-plot.

#### 2.4.4 Rotação ortogonal de fatores

Geralmente a interpretação dos fatores obtidos na Análise Fatorial pode não ser uma tarefa fácil, devido a ocorrência de coeficientes  $\hat{\lambda}_{ij}$  de grandeza numérica similar e não desprezível em fatores distintos. Quando isto ocorre, utiliza-se uma transformação ortogonal dos fatores originais, na tentativa de encontrar uma estrutura mais simples para ser interpretada. Para isso, utiliza-se uma matriz ortogonal  $T$  que pode ser obtida por diferentes critérios, sendo que após a rotação dos fatores, a estrutura de variância é preservada (Rencher, 2002). Segue adiante três critérios de rotação: varimax, orthomax e quartimax. Neste trabalho foram testados os três critérios, porém o critério varimax produziu melhores resultados, e é o único utilizado nas interpretações de cada Análise Fatorial aplicada ao conjunto de dados.

##### 2.4.4.1 Critério Varimax

De acordo com Mingoti (2005) o critério de rotação varimax é um método proposto por Kaiser (1958), cujo objetivo principal é a busca da matriz  $T$ , tendo como base encontrar para um determinado fator fixo, um grupo de variáveis  $Z_i$  altamente correlacionadas com o fator e outro grupo que tenha correlação desprezível ou moderada com o fator. Para cada fator fixo, a solução é obtida através da maximização da variação dos quadrados das cargas fatoriais originais das colunas da matriz  $\hat{L}$ .

Seja

$$V = \frac{1}{p} \sum_{j=1}^m \left[ \sum_{i=1}^p \tilde{l}_{ij}^4 - \frac{1}{p} \left( \sum_{i=1}^p \tilde{l}_{ij}^2 \right)^2 \right] \quad (2.110)$$

onde  $\tilde{l}_{ij} = \frac{\hat{l}_{ij}}{\hat{h}_i}$ , sendo  $\hat{h}_i$  a raiz quadrada da comunalidade da variável  $Z_i, i=1,2,\dots,p$ .

O critério varimax seleciona os coeficientes  $\tilde{l}_{ij}$  que maximizam  $V$ . Os coeficientes finais da matriz transformada são obtidos por:  $\hat{L}^* = \hat{L}T$ . O critério varimax é um dos mais utilizados na prática e, em geral, produz resultados mais simples que outros métodos.

#### 2.4.4.2 Critério Quartimax

Esse método busca a simplificação das linhas de uma matriz de cargas fatoriais. Segundo Hair et al. (2009) o critério quartimax de rotação deixa muitas variáveis com cargas fatoriais altas em um único fator, não sendo bem sucedido na produção de estruturas mais simples.

Seja  $V_Q$  a quantidade definida por:

$$V_Q = \frac{1}{pm} \left[ \sum_{j=1}^m \sum_{i=1}^p \hat{l}_{ij}^{*4} - \frac{1}{pm} \left( \sum_{j=1}^m \sum_{i=1}^p \hat{l}_{ij}^{*2} \right)^2 \right] \quad (2.111)$$

onde  $\hat{l}_{ij}^*$  é o coeficiente da  $i$ -ésima variável no  $j$ -ésimo fator após a rotação. O critério quartimax seleciona os coeficientes  $\hat{l}_{ij}^*$  que maximizam a quantidade  $V_Q$  (MINGOTI, 2005).



#### 2.4.4.3 Critério orthomax

É considerado uma média ponderada dos métodos varimax e quartimax (MINGOTI, 2005). Considere

$$V_M = \sum_{j=1}^m \left[ \sum_{i=1}^p \hat{l}_{ij}^{*4} - \frac{\gamma}{p} \left( \sum_{i=1}^m \hat{l}_{ij}^{*2} \right)^2 \right] \quad (2.112)$$

onde  $\hat{l}_{ij}^*$  é o coeficiente da  $i$ -ésima variável no  $j$ -ésimo fator após a rotação,  $0 \leq \gamma \leq 1$ .

O critério orthomax seleciona os coeficientes  $\hat{l}_{ij}^*$  que maximizam a quantidade  $V_M$ .

Para  $\gamma = 1$ , este critério reproduz o critério varimax sem o escalonamento das cargas fatoriais pelas comunalidades e, para  $\gamma = 0$ , ele se reduz ao método quartimax.

#### 2.4.5 Determinação dos escores fatoriais

Após a identificação e interpretação dos fatores  $L_j$ ,  $j=1,2,\dots,m$  relacionados com as variáveis  $Z_i$ ,  $i=1,2,\dots,p$  ( $i$ -ésima variável padronizada pela média e desvio padrão), é necessário calcular os escores para cada elemento amostral, de modo a utilizar esses valores para outras análises de interesse. Para cada elemento amostral  $k$ ,  $k=1,2,\dots,n$ , o seu escore no fator  $L_j$  é calculado por:

$$\hat{F}_{jk} = w_{j1}Z_{1k} + w_{j2}Z_{2k} + \dots + w_{jp}Z_{pk} \quad (2.113)$$

onde  $Z_{ik}$ , são valores observados das variáveis padronizadas  $Z_i$  para o  $k$ -ésimo elemento amostral e  $w_{ji}$ ,  $i=1,2,\dots,p$ , são os pesos de ponderação de cada variável  $Z_i$  no fator  $L_j$ . De acordo com Mingoti (2005), existem três métodos para a determinação dos coeficientes  $w_{ji}$ :

- ✓ método dos mínimos quadrados ponderados;
- ✓ método de regressão;
- ✓ método ad hoc.

Neste trabalho foram testados os três métodos, porem foi implementado apenas o método dos mínimos quadrados ponderados, uma vez que este apresentou um melhor resultados.

#### 2.4.5.1 Método dos mínimos quadrados ponderados

O modelo fatorial em (2.83) assemelha-se a um modelo de regressão linear múltipla. Desta maneira pode-se estimar o vetor  $\underline{F}$  através do método dos mínimos quadrados ponderados usando-se as matrizes  $\hat{L}$  e  $\hat{\Psi}$  estimadas pelo método da máxima verossimilhança. Sendo assim,  $\hat{F}_j$  pode ser obtido por:

$$\hat{F}_{jk} = (\hat{L}'\hat{\Psi}^{-1}\hat{L})^{-1}\hat{L}'\hat{\Psi}^{-1}Z_k = WZ_k \quad (2.114)$$

onde  $Z_k = (Z_{1k} \ Z_{2k} \ \dots \ Z_{pk})'$  é o vetor de observações do k-ésimo elemento amostral padronizado e  $W_{m \times p}$  é a matriz de ponderação que gera os coeficientes  $w_{ji}$ .

Neste trabalho, as matrizes  $\hat{L}$  e  $\hat{\Psi}$  foram estimadas pelo método das componentes principais. Neste caso, Johnson e Wichern (2007) sugere que a estimação dos escores seja feita usando o método dos mínimos quadrados ordinários, desde que os resíduos  $\varepsilon$  tenham a mesma variância, o que nem sempre ocorre (MINGOTI, 2005).

Desta forma,  $\hat{F}_j$  é estimado por:

$$\hat{F}_{jk} = (\hat{L}'\hat{L})^{-1}\hat{L}'Z_k = WZ_k \quad (2.115)$$

Os escores fatoriais finais ponderados servem para criar um indicador a partir de todos os escores fatoriais obtidos na Análise Fatorial. De acordo com Furtado et al. (2003) os escores fatoriais finais ponderados podem ser calculados da seguinte maneira:

$$\hat{E} = \frac{\sum_{i=1}^m \hat{\lambda}_i \cdot \hat{F}_i}{\sum_{i=1}^m \hat{\lambda}_i} \quad (2.116)$$

onde  $\hat{F}_i$  são os escores fatoriais do i-ésimo e  $m$  é o número de fatores extraídos.

## 2.5 LÓGICA NEBULOSA

A teoria dos sistemas nebulosos constitui uma metodologia particularmente adequada ao controle de sistemas, principalmente em situações com grau de complexidade elevada, onde a presença de não linearidades e de fatores de incerteza seja significativa. De fato, os problemas enunciados enquadram-se no conjunto das limitações marcantes das técnicas clássicas de controle e modelagem, pelo que os sistemas nebulosos se apresentam como um complemento importante das metodologias convencionais.

### 2.5.1 Teoria dos conjuntos nebulosos

Na natureza parte das propriedades são vagas, incertas ou imprecisas e, portanto, não são possíveis de serem caracterizadas por predicados da lógica clássica bivalente. A teoria dos conjuntos nebulosos pode ser vista como uma extensão da teoria clássica de conjuntos, e foi criada por Zadeh (1965) para tratar graus de pertinência intermediários entre a pertinência total e a não pertinência de elementos de um universo de discurso com relação a um dado conjunto. Assim, um conjunto nebuloso é uma generalização da noção clássica de um conjunto deste universo.

A definição de função de pertinência de um conjunto nebuloso depende do significado lingüístico definido para este conjunto e de sua interpretação no contexto do universo de discurso utilizado. Assim, na lógica nebulosa, o grau de pertinência de um elemento em relação a um conjunto é definido por uma função característica real, que assume como valor qualquer número pertencente ao intervalo real fechado  $[0, 1]$ . Algebricamente, é mapeada por  $\mu_y(x): U \rightarrow [0, 1]$ .

Então, um conjunto nebuloso  $F$  de um universo de discurso  $U$  é caracterizado por uma função característica real, comumente denominada função de pertinência, mapeada por  $\mu_F(x): U \rightarrow [0, 1]$ , que associa a cada  $x \in U$  um número real  $\mu_F(x)$  no intervalo  $[0, 1]$ , representando o grau de pertinência de  $x$  em  $F$ . O suporte de  $F$  é o subconjunto de pontos de  $U$  para o qual  $\mu_F(x)$  é positivo.

A representação de conjuntos nebulosos depende basicamente da natureza e da dimensão do universo de discurso definido. Quando o universo é discreto e composto de um número pequeno de elementos, a maneira mais usual adotada é a representação analítica.

Porém, quando o universo de discurso é contínuo ou possui uma quantidade grande de elementos discretos, uma forma simples de representação é o gráfico de sua função de pertinência, chamado de diagrama de Hassi-Euler (H-E).

Zadeh (1965) definiu o conceito de conjunto nebuloso como forma de superar a classificação dicotômica presente na teoria clássica dos conjuntos. Conseqüentemente, como resultado da aplicação dos conjuntos nebulosos aos mecanismos de inferência, o mesmo autor propôs vários algoritmos de processamento nebuloso para problemas específicos (ZADEH, 1973).

Do mesmo modo que os termos linguísticos presentes num algoritmo nebuloso devem ser quantificados por meio de conjuntos nebulosos, outras operações, tais como a interseção, união e implicação nebulosas são definidas por Zadeh (1973). A partir do momento em que os conjuntos nebulosos e os operadores estão determinados, as relações expressas qualitativamente pelas regras de controle comuns deixam de ser vagas, passando a constituir uma função não linear determinística. Obtém-se deste modo aquilo que se designa por sistema nebuloso: uma implementação específica, dependente do contexto de utilização, de um conjunto de regras qualitativas expressas por meio de um algoritmo nebuloso.

### 2.5.2 Variáveis linguísticas

Segundo Zadeh (1973), quanto mais a complexidade de um sistema aumenta, mais difícil de se produzir inferências precisas e ainda significantes sobre o seu comportamento. Sendo assim, deve-se utilizar outra maneira de se analisar o sistema. O raciocínio humano utiliza elementos no processo de tomada de decisão que não são números, mas sim, variáveis lingüísticas, cujo comportamento é nebuloso.

Uma variável linguística é definida como uma palavra  $x$  em uma linguagem natural  $L$ , a qual pode ser vista como uma descrição sumária de um subconjunto nebuloso  $M(x)$  em um universo de discurso  $U$ . Onde  $M(x)$  é o significado de  $x$ .

Por exemplo, supondo que  $x$  é uma variável como “Temperatura”, portanto podem-se definir os seguintes termos para essa variável:

$$M(x) = \{fria, boa, quente\} \quad (2.117)$$

onde  $U(x) = [20, 60]$ .

Portanto, uma possível função de pertinência dessa variável é apresentada na Figura 2.10.

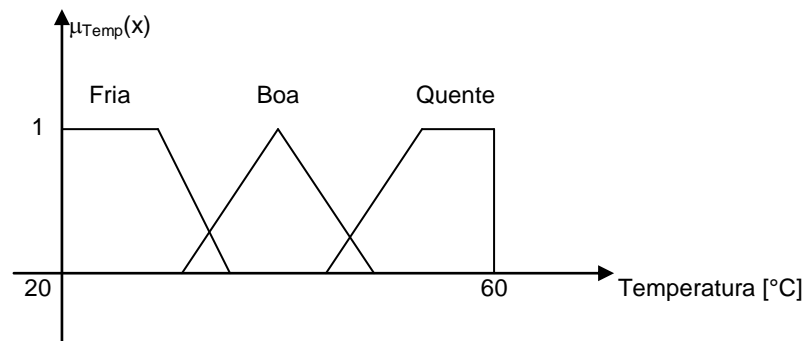


FIGURA 2.10 – CONJUNTO NEBULOSO DA VARIÁVEL “TEMPERATURA”  
 FONTE: O AUTOR

### 2.5.3 Controle nebuloso

LEE (1990) apresenta uma revisão sobre os desenvolvimentos pertinentes a controladores nebulosos, os quais compreendem quatro principais componentes apresentados na Figura 2.11, tal que:

- ✓ A interface de nebulização recebe os valores das variáveis de entrada e desenvolve uma conversão para uma variável linguística em conjunto com um grau de pertinência;
- ✓ A base de conhecimento contempla as regras de controle em termos de variáveis linguísticas;
- ✓ A lógica de tomada de decisão, a qual é capaz de emular o processo de tomada de decisão humano e inferir ações de controle nebulosas;

- ✓ A interface de desnebulização recebe as ações de controle nebulosas, em termos de uma variável lingüística em conjunto com um grau de pertinência e converte para valores das variáveis de saída.

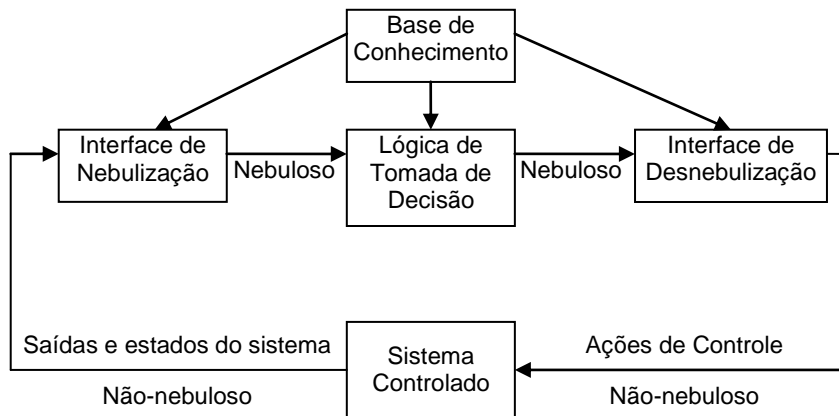


FIGURA 2.11 – CONFIGURAÇÃO BÁSICA DE UM CONTROLADOR NEBULOSO.  
 FONTE: O AUTOR

#### 2.5.4 Modelo Linguístico do Tipo Mamdani

Mamdani (1977) apresenta um método de inferência nebulosa que foi por muitos anos um padrão para a utilização dos conceitos da lógica nebulosa em processamento de conhecimento. As regras de produção em um modelo de Mamdani possuem relações nebulosas tanto em seus antecedentes como em seus conseqüentes, conforme apresentado em (2.118).

$$\mathbf{Se} \langle \text{antecedente} \rangle \mathbf{ent\~{a}o} \langle \text{consequente} \rangle \quad (2.118)$$

O modelo de Mamdani inclui módulos de interface que transformam as variáveis de entrada em conjuntos nebulosos e, posteriormente, os conjuntos nebulosos gerados na saída em grandezas numéricas proporcionais. A Figura 2.12 mostra um diagrama típico de um sistema de processamento de conhecimento nebuloso que utiliza o modelo de inferência Mamdani.

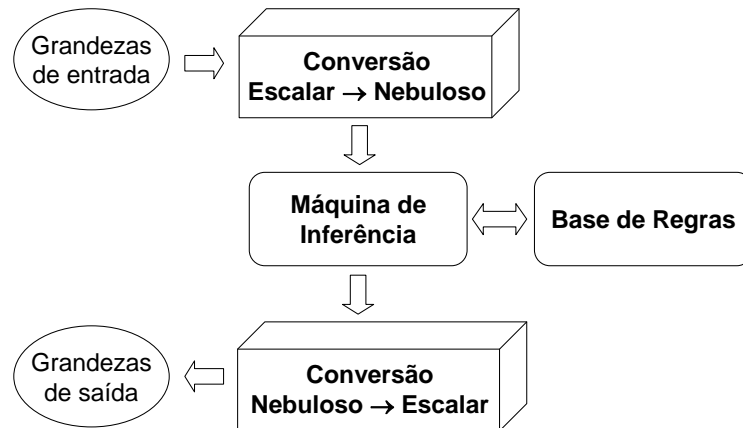


FIGURA 2.12 – DIAGRAMA TÍPICO DE UM MODELO DE INFERÊNCIA MAMDANI  
 FONTE: AOKI, 2005.

A regra semântica tradicionalmente utilizada para o processamento de inferências com o modelo de Mamdani é chamada de inferência Máx-Mín. Esta regra utiliza as operações de união e interseção entre conjuntos da mesma forma que Zadeh (1973), por meio dos operadores de máximo e mínimo, respectivamente. Considerando uma regra de produção nebulosa genérica mostrada em (2.119), o método de inferência Máx-Mín proposto por Mamdani será mostrado em seguida.

$$\text{Se } x_1 = A_i \text{ e } x_2 = A_j \text{ e } \dots \text{ e } x_p = A_k \text{ então } y_1 = B_i \text{ e } y_2 = B_m \quad (2.119)$$

onde:

$x_i$  são as entradas do sistema;

$A_i$  são os termos primários definidos nos conjuntos nebulosos;

$y_i$  são as saídas do sistema;

$B_i$  são os termos primários definidos nos conjuntos nebulosos.

Durante o processo de conversão escalar para nebuloso (nebulização), os antecedentes de cada regra são processados por meio de intersecção nebulosa (minimização) entre os graus de pertinência das entradas atuais nos termos primários definidos em cada uma. Este processo gera um grau de pertinência de disparo para cada regra de produção. Em outras palavras, é calculado para a  $k$ -ésima regra da base de conhecimento um coeficiente de disparo  $D(k)$  segundo (2.120), onde os índices  $k$  nos conjuntos nebulosos denotam os termos primários que compõem a regra  $k$  na base de conhecimento. Este processamento transforma informações quantitativas em informações qualitativas e é considerado um processo de generalização, tal que:

$$D^k = T[\mu_{A_1^k}(x_1), \mu_{A_2^k}(x_2), \dots, \mu_{A_p^k}(x_p)] = \min[\mu_{A_1^k}(x_1), \mu_{A_2^k}(x_2), \dots, \mu_{A_p^k}(x_p)] \quad (2.120)$$

As regras para as quais o coeficiente de disparo for maior que zero são ditas regras que dispararam para as entradas atuais. Isto quer dizer que elas vão contribuir para o cálculo de saída correspondente do sistema de inferência. Os coeficientes de disparo, por sua vez, vão limitar os valores máximos dos conjuntos nebulosos de saída gerados por estas regras.

Finalmente, uma operação global de união vai compor um conjunto nebuloso para cada variável de saída, contendo informações sobre todas as regras disparadas para as entradas atuais. Em (2.121) é mostrada a composição deste conjunto para o caso da saída  $y_2$  da regra, mostrada em (2.119). O universo de discurso desta variável de saída é composto pelos elementos  $y \in U_{y_2}$ , tal que

$$\mu_{B_i}(y) = S_{k=1..n} [T(D^{(k)}, \mu_{B_i}(y))] = \max_{k=1..n} [\min(D^{(k)}, \mu_{B_i}(y))] \quad \forall y \in U_{y_2} \quad (2.121)$$

O processo de inferência acima descrito transforma uma informação qualitativa em outra informação qualitativa, por meio de uma conversão. O conjunto nebuloso gerado durante o processo de inferência pode então ser utilizado diretamente em um diagnóstico qualitativo de tomada de decisão, ou será convertido em um valor escalar proporcional para atuação externa por meio de atuadores convencionais, no caso de um sistema de controle.

A conversão nebulosa para escalar (desnebulização) transforma informações qualitativas em uma informação quantitativa, sendo um processo de especificação. Para a regra de produção definida em (2.119), parte-se de um conjunto nebuloso de saída  $\mu_{B_i}(y)$  obtido pelo processo de inferência e chega-se a um valor escalar  $\hat{y}_2$  correspondente. Os métodos mais utilizados para a conversão nebulosa para escalar são o método do centro de massa e o método da média dos máximos.



### 3 MATERIAL E MÉTODO

#### 3.1 MATERIAL

No desenvolvimento deste trabalho foram utilizados dois bancos de dados. O primeiro banco de dados contempla as medições de qualidade da energia elétrica feitas no secundário de 359 transformadores da rede de distribuição da Companhia Paranaense de Energia (COPEL). O segundo banco de dados composto de características físicas e elétricas de 181 estruturas metálicas de linhas de transmissão (LTs) da Companhia Paranaense de Energia (COPEL). As descrições de todos os conjuntos de dados são detalhadas a seguir:

##### 3.1.1 Caso I – Campanha de Medição de Distorções Harmônicas de Tensão

A Copel atende cerca de 100% das residências urbanas no Paraná e 90 % das rurais, possuindo no total cerca de 3,39 milhões de consumidores ligados à sua rede (COPEL, 2007).

Do total de consumidores, 2,6 milhões são residências, 57 mil indústrias, 283 mil estabelecimentos comerciais e 332 mil propriedades rurais e a cada ano, em média, são realizadas 70 mil novas ligações.

As Tabelas 3.1 e 3.2 descrevem algumas características adicionais referentes à Copel:

TABELA 3.1 – CARACTERÍSTICAS DA COPEL PARA A ATUAÇÃO NO PARANÁ

<b>Indicadores</b>	<b>Valores</b>
Áreas de concessão (km <sup>2</sup> )	194854
Municípios atendidos	393
Localidades atendidas	1111
População atendida PR (Mil Habitantes)	9822
Taxa de atendimento urbano (%)	99
Consumo médio residencial (kWh/mês)	159,6

FONTE: (COPEL, 2007)

TABELA 3.2– CARACTERÍSTICAS DA COPEL – DADOS DA DISTRIBUIÇÃO

<b>Dados de Distribuição</b>	<b>Valores</b>
Número de subestações 34,5 kV	238
Potência instalada (MVA)	1576
Redes de distribuição (km)	167535
Total de postes	2.296301
<b>Número de transformadores de distribuição</b>	<b>315587</b>
Potência instalada em transformadores (MVA)	8009

FONTE: (COPEL, 2007)

A estrutura da Copel no estado do Paraná está dividida em cinco superintendências sendo que cada uma delas atua em uma região do estado e são assim denominadas:

- Superintendência de Distribuição Leste – SDL, com sede em Curitiba;
- Superintendência de Distribuição Oeste – SDO, com sede em Cascavel;
- Superintendência de Distribuição Centro-Sul – SDC, com sede em Ponta Grossa;
- Superintendência de Distribuição Noroeste – SDN, com sede em Maringá;
- Superintendência de Distribuição Norte – SDT, com sede em Londrina.

A região que possui o maior consumo no estado é a região leste (SDL-35,6%), seguida da região centro sul (SDC-18,6%), noroeste (SDN-17,3%), norte (SDT-14,8%) e região oeste (SDO-13,7%).

### 3.1.1.1 Medições

Devido ao fato de que existem muitos transformadores na rede de distribuição da Copel no estado do Paraná, a medição de todos eles se torna uma prática inviável. Para solucionar este problema buscou-se obter uma amostragem que preservasse as informações do universo estudado, uma das ações seguidas foi a tomada de amostra aleatória dos transformadores que seriam medidos.

Outra ação foi a obtenção do número de amostras que preservariam as informações do universo estudado, para isso, foi consultada a portaria 163/93 do

Departamento Nacional de Águas e Energia Elétrica (DNAEE) e obteve-se uma fórmula específica para este fim, conforme equação 6, tal que:

$$C_a = \frac{N.S.\bar{S}}{\frac{(N-1).E^2}{4} + (S.\bar{S})} \quad (3.1)$$

onde

- $C_a$  - tamanho da amostra;
- $N$  - tamanho do universo considerado = 315587;
- $S$  - proporção da população com características consideradas semelhantes, considerado como 0,5 (pior caso);
- $\bar{S}$  - proporção da população com características não semelhantes àquelas de  $S$ , considerado como 0,5 (1- $S$ );
- $E$  - erro amostral = 5%.

Portanto, aplicando-se a (equação 3.1), de um universo total de 315.587 registros, chegou-se a um tamanho de amostra de 399 transformadores a serem monitorados na campanha de medição. Contudo, prevendo-se as dificuldades que poderiam ser encontradas durante as medições, tais como deslocamento, impossibilidade de medição, qualidade das medidas a serem registradas, perda de dados, entre outras foi estipulada uma margem de segurança de 10%, com isso, o número definitivo de registros a serem observados foi de 440 transformadores.

TABELA 3.3 – DISTRIBUIÇÃO DOS PONTOS DE MONITORAÇÃO

Regional	Localização	Região	Amostras (%)	Amostras
SDC	Centro	Ponta Grossa	11%	49
SDL	Leste	Curitiba	33%	146
SDN	Noroeste	Maringá	18%	77
SDO	Oeste	Cascavel	19%	84
SDT	Norte	Londrina	19%	84
Total			100%	440

FONTE: (LACTEC, 2007)

TABELA 3.4 – DISTRIBUIÇÃO DAS AMOSTRAS ENTRE ÁREAS URBANAS E RURAIS

<b>Regional</b>	<b>Amostras Urbanas</b>	<b>Amostras Rurais</b>
SDC	38	11
SDL	136	10
SDN	68	9
SDO	69	15
SDT	76	8
Total	387	53

FONTE: (LACTEC, 2007)

Na tabela 3.3 é possível visualizar a distribuição dos pontos entre as cinco superintendências e na tabela 3.4, a distribuição dos pontos entre transformadores rurais e urbanos.

O tempo utilizado para a realização das medições foi de 14 meses, no período de agosto de 2005 a julho de 2007. Foram utilizados 12 aparelhos analisadores de energia, registrador MARH 21, descrito a seguir.

- Fabricante: RMS Sistemas Eletrônicos – Brasil;
- 128 amostras por ciclo;
- Período de integralização de 10 minutos;
- Memória de 4MB;
- Tensão alimentação: CA 70 a 600 V, frequência: 50Hz ou 60Hz;
- CC 11 a 600 V;
- Exatidão: classe 0,5 para tensões e correntes e classe 1 para potências.

As medições foram realizadas com os medidores instalados por um período de oito dias em média para cada transformador, porém, para a análise e composição do banco de dados utilizado neste trabalho foram considerados os dados obtidos em um período de sete dias, devido ao disposto no Submódulo 2.2 do ONS e da Resolução n. 505 da ANEEL, que se referem aos intervalos para análise. Na figura 3.1, a seguir, está ilustrado como exemplo o período de monitoração e o período de análise utilizado em uma amostra (transformador).

Os valores obtidos nas medições de DHTV foram dispostos conforme as figuras 3.2 e 3.3, onde é possível, no eixo vertical, obter o valor da distorção em relação à fundamental, em cada fase. Estas medições foram realizadas no barramento de 127/220V.

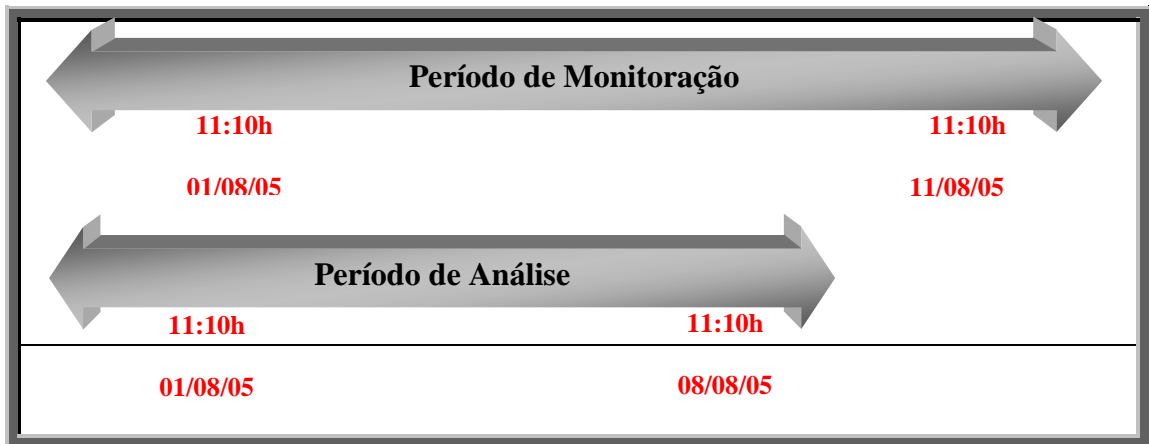


FIGURA 3.1 – PERÍODO DE MONITORAÇÃO E PERÍODO DE ANÁLISE  
 FONTE: (LACTEC, 2007)

Nas figuras 3.2 e 3.3 tem-se que:

- DHTVa - níveis de distorção harmônica total de tensão na fase A;
- DHTVb - níveis de distorção harmônica total de tensão na fase B;
- DHTVc - níveis de distorção harmônica total de tensão na fase C.

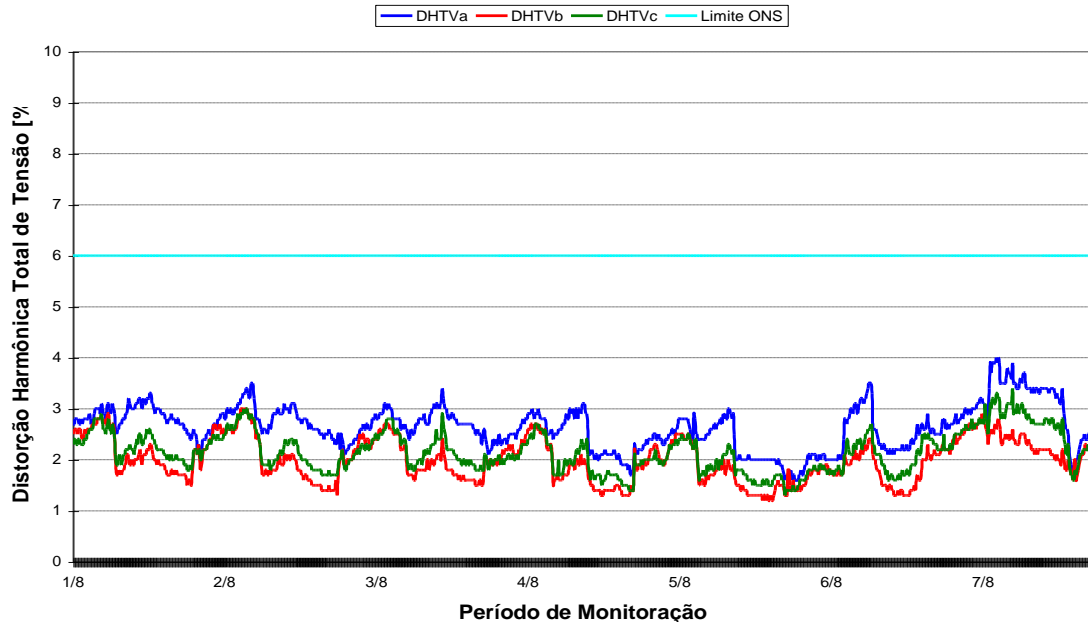


FIGURA 3.2 – EXEMPLO DE DHTV – ABAIXO DE 6%  
 FONTE: (LACTEC, 2007)

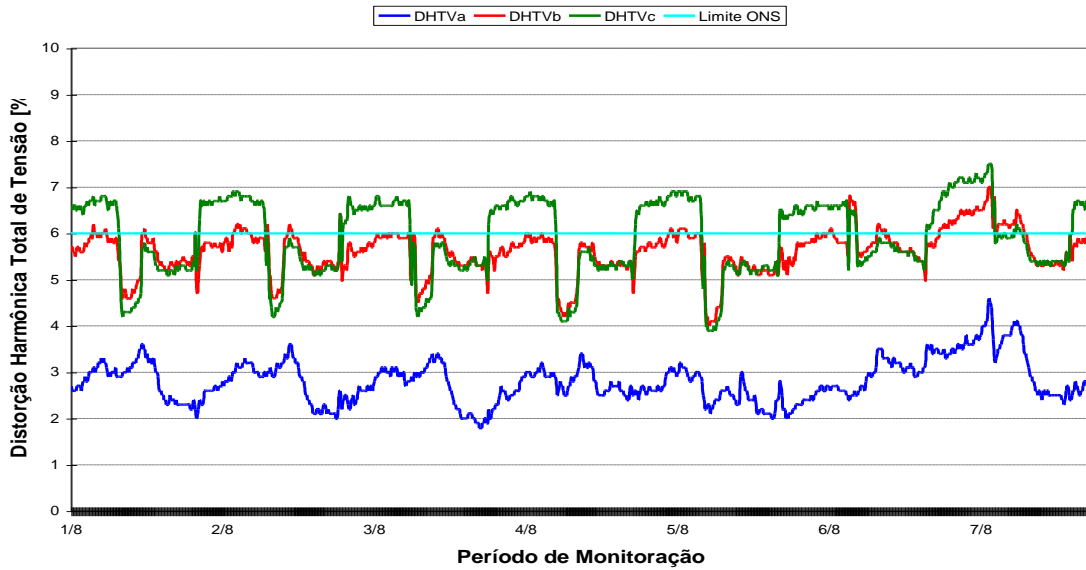


FIGURA 3.3 – EXEMPLO DE DHTV – ACIMA DE 6%  
 FONTE: (LACTEC, 2007)

Algumas informações importantes podem ser extraídas das medições da distorção Harmônica Total de Tensão, tais como valores máximos e mínimos obtidos durante o monitoramento. As tabelas 3.5 e 3.6 a seguir demonstram exemplos retirados destas informações extraídas das figuras 3.2 e 3.3.

TABELA 3.5 – VALORES MÁXIMOS, MÍNIMOS E P<sub>95%</sub> DA DHTV - DENTRO DO LIMITE (6%)

DHV <sub>T</sub>	DHTV <sub>A</sub> [%]	DHTV <sub>B</sub> [%]	DHTV <sub>C</sub> [%]
<b>Máximo</b>	4,70	4,90	4,60
<b>Mínimo</b>	2,30	2,20	2,40
<b>P<sub>95%</sub></b>	4,53	4,66	4,35

FONTE: (LACTEC, 2007)

TABELA 3.6 – VALORES MÁXIMOS, MÍNIMOS E P<sub>95%</sub> DA DHTV - FORA DO LIMITE (6%)

DHV <sub>T</sub>	DHTV <sub>A</sub> [%]	DHTV <sub>B</sub> [%]	DHTV <sub>C</sub> [%]
<b>Máximo</b>	4,60	7,00	7,50
<b>Mínimo</b>	1,80	4,00	3,90
<b>P<sub>95%</sub></b>	4,09	6,70	7,30

FONTE: (LACTEC, 2007)

Durante o monitoramento foi possível obter também o espectro harmônico, ou seja, o valor em porcentagem de cada frequência harmônica de tensão. Nas figuras 3.4 e 3.5 são ilustradas estas medições para o período de maior distorção conforme exemplos citados nas figuras 3.2 e 3.3 respectivamente.

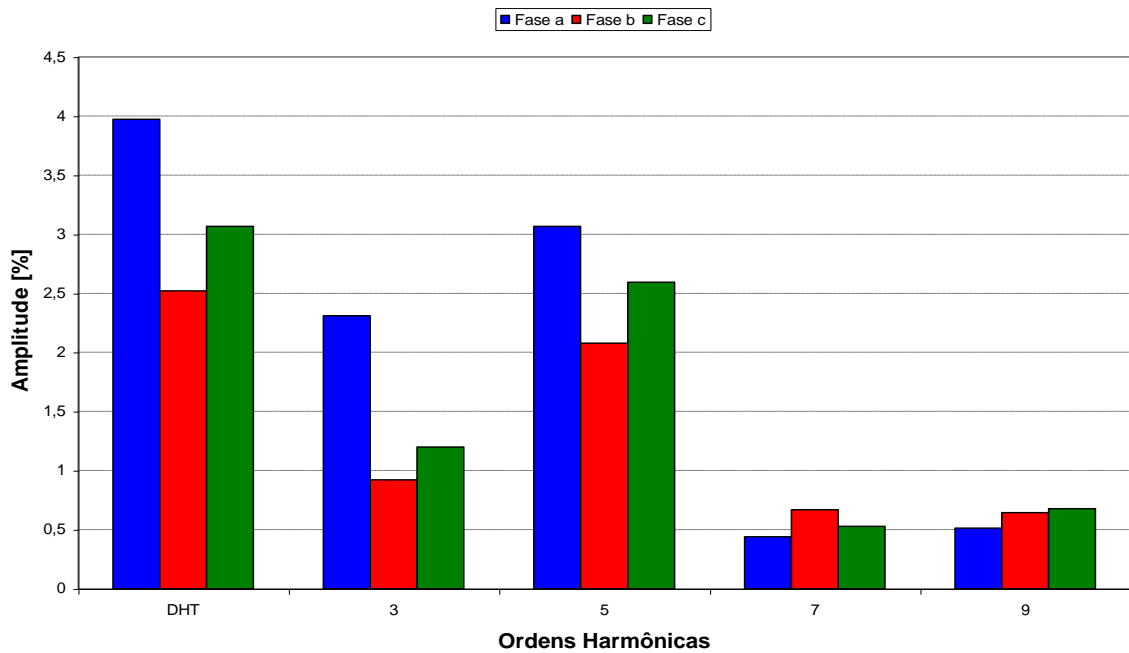


FIGURA 3.4 – ESPECTRO DA TENSÃO NO PERÍODO DE MAIOR DISTORÇÃO – ABAIXO DE 6%  
 FONTE: (LACTEC, 2007)

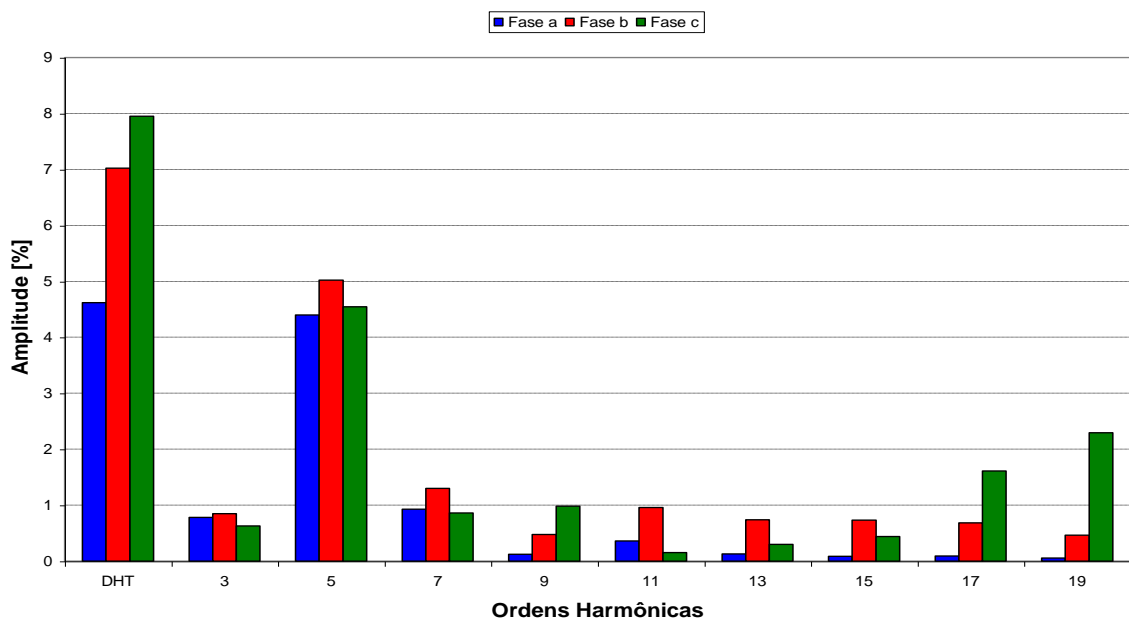


FIGURA 3.5 – ESPECTRO DA TENSÃO NO PERÍODO DE MAIOR DISTORÇÃO – ACIMA DE 6%  
 FONTE: (LACTEC, 2007)

Durante o monitoramento dos transformadores, foram encontradas diversas dificuldades já previstas, tais como: dificuldade de deslocamento, impossibilidade de medição devido à alguma impossibilidade física, etc. Com isso, ao final da

campanha de medições, foram medidos no total 359 transformadores urbanos distribuídos em 186 cidades do estado do Paraná, conforme exposto na tabela 3.7.

TABELA 3.7 – NÚMERO DE TRANSFORMADORES URBANOS MEDIDOS

<b>Superintendência</b>	<b>SDL</b>	<b>SDT</b>	<b>SDO</b>	<b>SDC</b>	<b>SDN</b>	<b>TOTAL</b>
Número de cidades visitadas	29	42	55	23	37	186
Quantidade de transformadores	128	74	62	32	63	359

FONTE: (LACTEC, 2007)

### 3.1.1.2 Variáveis Utilizadas

O estudo obedeceu aos requisitos do Submódulo 2.2 do ONS (2002), que estabelece como recomendação quanto aos índices sobre distorções harmônicas totais e individuais de tensão que o máximo permitido para a distorção harmônica total de tensão em qualquer fase seja o percentil 95%, ou seja, ( $P_{95\%}$ ). A variável dependente considerada no estudo foi o  $P_{95\%}$  de distorção harmônica total acima de 6% e abaixo de 6%, ou seja, os grupos foram separados tendo às seguintes características:

- Grupo 0 tendo como variável dependente o  $P_{95\%}$  da distorção harmônica total de tensão menor que 6% em todas as fases do transformador;
- Grupo 1 tendo como variável dependente o  $P_{95\%}$  da distorção harmônica total de tensão maior que 6% em qualquer uma das fases do transformador;

As variáveis independentes ou explicativas foram separadas em características do transformador e características de consumo. As variáveis correspondentes as características dos transformadores estão apresentadas na tabela 3.8 e as variáveis correspondentes as características de consumo dos transformadores estão apresentadas na tabela 3.9.



TABELA 3.8 – VARIÁVEIS ASSOCIADAS À CARACTERÍSTICA DO TRANSFORMADOR

<b>Características do Transformador</b>	<b>Classes</b>
Superintendência Regional	• SDL
	• SDN
	• SDO
	• SDC
	• SDT
Tensão Primária (Nível de Tensão do Alimentador)	• 13.800V
	• 34.500V
Fases (Quantidade de Fases no Transformador)	• 1 Fase
	• 2 Fases
	• 3 Fases
Potência (Potência Nominal do Transformador (kVA))	• 112,5 kVA
	• 75 kVA
	• 45 kVA
	• 30 kVA
	• 15 kVA
	• 10 kVA
	• 5 kVA

FONTE: COPEL 2007

TABELA 3.9 – VARIÁVEIS ASSOCIADAS À CARACTERÍSTICA DO CONSUMO.

<b>Características do Consumo</b>	
Consumo Médio Mensal	Número de Consumidores
	Residencial 0/50 kWh
	Residencial 50/100 kWh
	Residencial 100/150 kWh
	Residencial 150/200 kWh
	Residencial 200/300 kWh
	Residencial 300/500 kWh
	Residencial 500/1000 kWh
	Residencial >1000 kWh
	Comércio e Poder Público
	Industrial
	Rural
	Iluminação Pública

FONTE: COPEL 2007

### 3.1.2 Caso II – Histórico de Quedas de Estruturas de LTs no Estado do Paraná

Para realizar as análises com das quedas de Estruturas, foi necessário fazer uma pesquisa de informações em relação às estruturas que caíram, a partir dos dados históricos de quedas de estruturas metálicas.

Estes dados foram obtidos da COPEL, que além dos dados tabelados, possui relatórios impressos e manuscritos sobre as ocorrências registrados e arquivados para sua consulta, quando necessário.

Partindo dos atributos necessários, foi feita uma pesquisa sobre os relatórios antigos para este levantamento de dados. Além disto, foi feito um rastreamento para identificar as condições em que as estruturas estavam quando caíram. As informações técnicas das estruturas foram identificadas nos relatórios, enquanto os atributos de ordem climática ou relacionada à sua posição geográfica foram levantadas pelo posicionamento aproximado em relação às estruturas que foram implantadas em substituição das danificadas.

Até 2010, a rede elétrica do estado do Paraná registrou 61 quedas de estruturas ao longo dos últimos 50 anos de operação e manutenção de linhas de transmissão, sejam de 69 até 525 kV. Destas, mais de 40 linhas de transmissão diferentes foram afetadas e aproximadamente 150 torres foram reconstruídas.



FIGURA 3.6 – QUEDA DA ESTRUTURA 93 DA LT IRATI – SABARÁ  
FONTE: WAZEN 2011



FIGURA 3.7 – QUEDA DA ESTRUTURA 34 DA LT MEDIANEIRA – CÉU AZUL  
 FONTE: WAZEN 2011

A figura 3.6 ilustra a queda da estrutura 93 ocorrida na linha de 138 kV que interliga a Subestação Irati à Sabará (em Ponta Grossa), enquanto que a figura 3.7 apresenta a queda da estrutura 34 da linha de transmissão 138 kV que interliga a subestação Medianeira à Céu Azul.

A Figura 3.8 apresenta todas as linhas de transmissão que possuem pelo menos uma estrutura que sofreu queda e teve a mesma reconstruída no estado do Paraná.

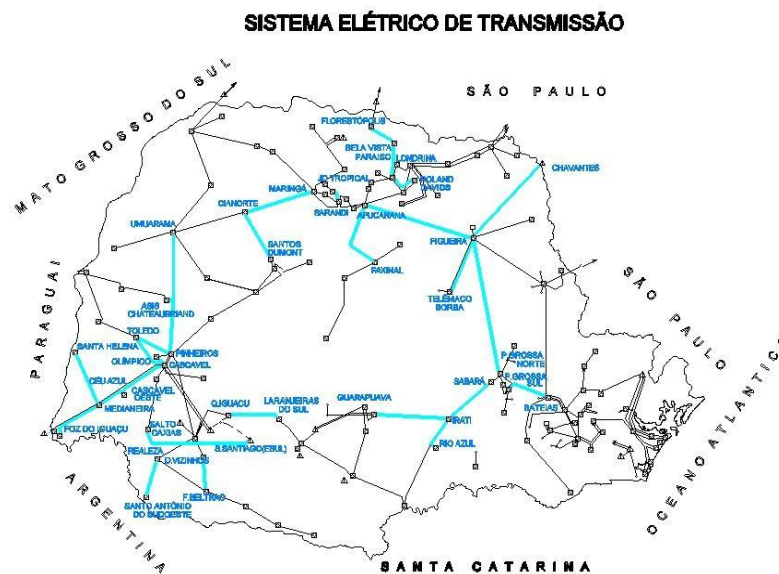


FIGURA 3.8 – IDENTIFICAÇÃO DAS LINHAS ONDE OCORRERAM QUEDAS  
 FONTE: COPEL



ocorrências. A partir destas informações e buscando informações que tenham ficado ocultas em um primeiro instante, foram definidos quais atributos podem interferir na queda de uma estrutura metálica.

Portanto, os atributos selecionados para a aplicação são: tensão de operação, vento na região, relevo do local, massas de ar, temperatura da região, tipo de terreno, capacidade mecânica da estrutura, função e tipo de fundação da estrutura. A tabela 3.10 apresenta os atributos selecionados bem como a variável numérica associada a cada atributo. Os grupos foram separados tendo às seguintes características:

- Grupo 0 sendo o grupo de estruturas que não apresentou queda;
- Grupo 1 sendo o grupo de estruturas que apresentou queda.

TABELA 3.10 – RELAÇÃO DE CARACTERÍSTICAS NÃO-NUMÉRICAS PARA NUMÉRICAS

Característica	Variável Não-numérica	Variável Numérica
Relevo	Planalto	1
	Planície	2
	Serra	3
	Vale	4
Massa de Ar	Paralela	1
	Perpendicular	2
	Transversal	3
Terreno	A	1
	B	2
	C	3
	D	4
Capacidade	Alta	1
	Baixa	2
Função	Ancoragem	1
	Suspensão	2
Fundação	Grelha	1
	Stub	2
Resultado	Nenhum	1
	Queda	0

FONTE: WAZEN 2011

## 3.2 MÉTODO

### 3.2.1 Introdução

Durante as últimas décadas houve um considerável crescimento do interesse em problemas de Reconhecimento e Classificação de Padrões, com consequente aumento da necessidade de métodos para uso na elaboração de sistemas nesta área.

Na engenharia de potência existem diversos trabalhos que contemplam métodos de Reconhecimento e Classificação de Padrões, em muitos trabalhos são feitas comparações de métodos com o intuito de encontrar um melhor método para o problema estudado.

Souza (2008) desenvolvimento de um modelo estatístico computacional para a estimação das distorções harmônicas de tensão a partir das características dos transformadores de distribuição e dos consumidores utilizando-se de técnicas de Análise Multivariada.

Martins (2010) utilizou o mesmo banco de dados utilizados por Souza (2008). O objetivo final do trabalho foi o mesmo, porém foi utilizada a técnica de *rough sets*. Os resultados obtidos foram comparados ao trabalho apresentado por Souza (2008).

Wazen (2011) avaliou os métodos de Reconhecimento e Classificação de Padrões de Teoria de Conjuntos Aproximados (Rough Sets) e Regressão Logística em um conjunto de dados que contempla quedas de estruturas metálicas na rede de distribuição.

Com base no cenário descrito, observamos que uma arquitetura Multiagentes que integre as vantagens e desempenho de diversos métodos de reconhecimento e classificação de padrões em um único sistema se torna atraente. A arquitetura Multiagentes foi desenvolvida com o objetivo de gerenciar métodos de Reconhecimento e Classificação de Padrões e Mineração de Dados das variáveis de entrada dos modelos bem como gerenciar a combinação dos resultados dos métodos aplicados.

### 3.2.2 Sistema Multiagentes

Segundo Aoki (2003), um Sistema Multiagente (SMA) tem como premissa básica prover mecanismos para a criação de sistemas inteligentes a partir de entidades de software autônomas, chamadas de agentes. Esses agentes interagem através de um ambiente, compartilhado por todos os membros do SMA, e sobre o qual eles atuam alterando o seu estado.

Uma vez que cada Agente Inteligente possui certa habilidade e objetivos próprios, o Sistema Multiagentes tem que ser capaz de prover mecanismos para interação e coordenação dessas entidades. Dessa forma, o Sistema Multiagentes consegue resolver problemas maiores que os domínios de conhecimento de seus integrantes.

Os Sistemas Multiagentes podem ser classificados como reativos e cognitivos. Os sistemas reativos são compostos por um grande número de agentes muito simples que interagem entre si, sem possuir uma representação explícita do estado do ambiente e dos outros agentes, nem de suas ações passadas (FERBER, 1999).

Já os sistemas cognitivos em geral possuem poucos agentes, dado que cada agente comporta-se de maneira racional, possui uma concepção lógica simbólica do mundo e planeja suas estratégias de ação de acordo com um forte mecanismo de inferência lógica. Esse tipo de agente, além de utilizar técnicas de dedução e aprendizado, considera também aspectos característicos da vontade humana, como crença, desejo e intenção (WOOLDRIDGE & JENNINGS, 1998).

Segundo Gasser e Bond (1988) para se obter um comportamento coerente em um Sistema Multiagentes, os agentes deveriam não somente compartilhar conhecimento sobre o problema e as soluções, mas também raciocinar sobre o processo de coordenação entre os agentes.

#### 3.2.2.1 Arquitetura do Sistema Multiagentes Proposta

No desenvolvimento da Arquitetura do Sistema Multiagentes foi utilizado o software desenvolvido pela *Java Agent Development Framework* - JADE, que é um framework de software totalmente implementado em linguagem Java, que tem por



objetivos simplifica a implementação de Sistemas Multiagentes através de um *middle-ware* que está em conformidade com as especificações *Foundation for Intelligent Physical Agents - FIPA* e através de um conjunto de ferramentas gráficas que suporta a depuração e fases de implantação.

Esta plataforma de agentes pode ser distribuída através de máquinas e a configuração pode ser controlada através de uma interface gráfica remota. A configuração pode ser alterada, mesmo em tempo de execução, movendo agentes de uma máquina para outra, como e quando necessário.

Todas as operações entre os agentes são feitas dentro da arquitetura desenvolvida no JADE, porém toda operação feita dentro de cada agente é processada em MATLAB R2010b.

Para o desenvolvimento de um Sistema Multiagentes é necessário observar que o sistema é composto por várias entidades que irão interagir entre si para alcançar um objetivo global, o primeiro passo é identificar o objetivo de cada agente individualmente, bem como o serviço por ele oferecido. Com isso deve-se identificar e classificar cada agente em categorias

A abordagem empregada na arquitetura proposta prevê duas categorias de agentes que são:

- Agentes Especializados: são detentores do conhecimento específico empregado na análise e solução dos problemas encontrados no domínio de aplicação.
- Agentes de Suporte: são os responsáveis por prestar serviços que auxiliem o processo de tomada de decisões, como: estabelecer e coordenar a comunicação entre os agentes, acessar banco de dados, interface com o mundo real, entre outros.

A arquitetura do Sistema Multiagentes proposta neste trabalho, apresentada na figura 3.10 é composta por 4 tipos de agentes: O Agente Minerados de Dados, o Agente de Classificação, o Agente de Comunicação e o Agente Decisor.



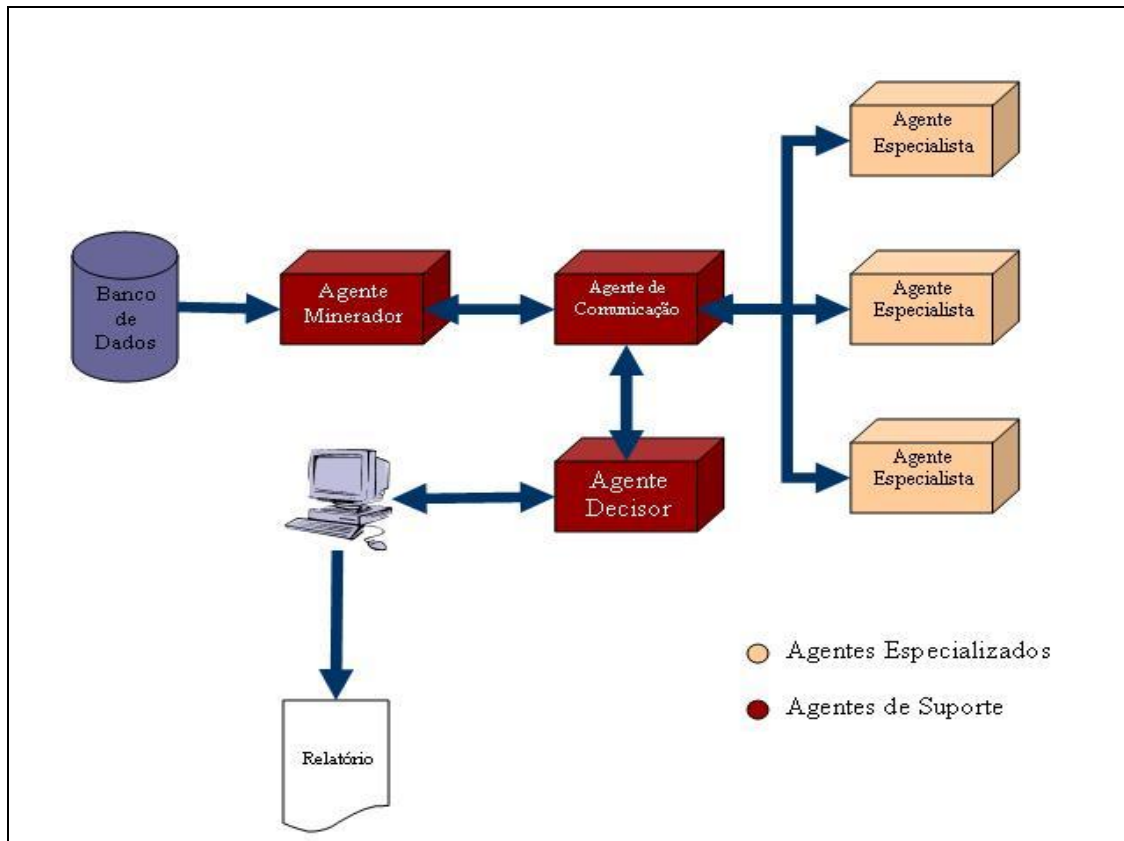


FIGURA 3.10 – PROPOSTA DE ARQUITETURA DO SISTEMA MULTIAGENTES  
 FONTE: O AUTOR

### 3.2.3 Agentes de Suporte

Os agentes de suporte, considerando que o sistema deve interagir com usuário e as informações utilizadas estarão contidas em algum tipo de base de dados num certo formato de apresentação prevê a necessidade de três agentes que são:

#### 3.2.3.1 Agente Minerador de Dados

O Agente Minerador de Dados é responsável por acessar a base de dados automatizada e realizar o pré-processamento da informação. Neste agente é descrita a metodologia de interpretação dos dados e filtragem de dados.

Segundo Fayyad, Piatetsky-Shapiro, & Smith (1996) a Mineração de Dados é definida como a extração de conhecimento da base de dados e é o processo de

identificação de padrões válidos, novos, potencialmente úteis e compreensíveis embutidos nos dados.

A Mineração de Dados é uma tecnologia usada para revelar informação estratégica escondida em grandes massas de dados. É usada em diversas áreas, como análise de riscos, marketing direcionado, controle de qualidade, análise de dados científicos, etc. Mineração de Dados define o processo automatizado de captura e análise de enormes conjuntos de dados, para então extrair um significado.

O processo de extração de conhecimento pode ser dividido em várias etapas, que são:

➤ Conhecimento e identificação do domínio do problema;

A identificação do problema contempla o estudo do domínio da aplicação e a definição de objetivos e metas a serem alcançadas no processo de Mineração de Dados.

➤ Pré-processamento da informação;

Normalmente, os dados disponíveis para análise não estão em um formato adequado para a extração de conhecimento. Além disso, em razão de limitações de memória ou tempo de processamento, muitas vezes não é possível a aplicação direta dos algoritmos de extração de padrões aos dados.

Dessa maneira, torna-se necessária a aplicação de métodos para tratamento, limpeza e redução do volume de dados antes de iniciar a etapa de extração de padrões. É importante observar que a execução de transformações na base de dados deve ser guiada pelos objetivos do processo de extração a fim de que o conjunto de dados gerados apresente as características necessárias para que os objetivos sejam alcançados.

➤ Extração de padrões;

A etapa de extração de padrões é direcionada ao cumprimento dos objetivos definidos na identificação do problema. Nessa etapa são realizadas a escolha, a configuração e a execução de um ou mais algoritmos para extração do conhecimento.

➤ Pós-processamento da informação;

Os algoritmos de extração de padrões podem gerar uma quantidade enorme de padrões. Porém, muitos dos quais podem não ser importantes, relevantes ou interessantes para o usuário. Portanto, é de vital importância desenvolver algumas técnicas de apoio no sentido de fornecer aos usuários apenas os padrões mais interessantes (Silberschatz & Tuzhilin 1995).

Um problema bastante comum que surge quando se utilizam técnicas de Mineração de Dados na etapa de pré-processamento é a necessidade de reduzir a dimensionalidade do problema estudado. O processo de redução é utilizado quando o número de variáveis estudadas é muito grande. Para reduzir a dimensão do vetor original pode-se criar um novo vetor cujas componentes são combinações lineares das variáveis originais. A redução da dimensionalidade do vetor de variáveis originais pode ser feita segundo dois critérios.

- Conhecimento informal do especialista utilizando a própria vivência em torno do problema, ou seja, conhecimento empírico.
- Conhecimento Científico através de critérios estatísticos, conhecidos como técnicas de seleção de variáveis.

A criação de um novo vetor de variáveis de menor dimensão, cujas componentes são combinações lineares das variáveis originais, pode ser conduzida via uma técnica de análise multivariada, conhecida como Análise de Componentes Principais .

Outro critério interessante e muito utilizado para reduzir a dimensionalidade dos problemas é a aplicação da Análise Fatorial, que utiliza a técnica de componentes principais para estimar os pesos fatoriais e possibilita a visualização de grupos que agregam variáveis originais que têm uma determinada associação e estão nas direções de maior variabilidade. Essas combinações lineares das variáveis originais são chamadas de fatores. A Análise Fatorial possibilita a avaliação da importância de cada variável original por meio do conceito de comunalidade.

Desta forma o Agente Minerador de Dados tem seu objetivo dividido em três partes principais.

- I. Ao acessar o banco de dados o Agente Minerador de Dados executa um pré-processamento dos dados, fazendo assim uma limpeza dos dados, retirando todas as redundâncias, seguidamente é aplicado os testes de Bartlett e KMO para verificar a validade da Análise Fatorial.
- II. A seguir o Agente Minerador de Dados utiliza a Análise Fatorial para extrair os fatores que explica a correlação existente entre as variáveis originais do banco de dados. Esses fatores agrupam as variáveis iniciais que contenham forte correlação, desta forma, as variáveis originais agrupadas num mesmo fator possuem, portanto alta correlação, enquanto que variáveis de fatores distintos possuem baixa correlação.
- III. Em posse dos autovalores calculados o Agente Minerador de Dados, utiliza o critério de Kaiser para a seleção dos fatores. A seguir com base na rotação Varimax e calculadas as comunalidades para cada variável, são calculados os escores fatoriais das observações que serão utilizados como a nova base de dados para os modelos de Reconhecimento e Classificação de Padrões.

A figura 3.11 apresenta o processo realizado pelo Agente Minerador de Dados.

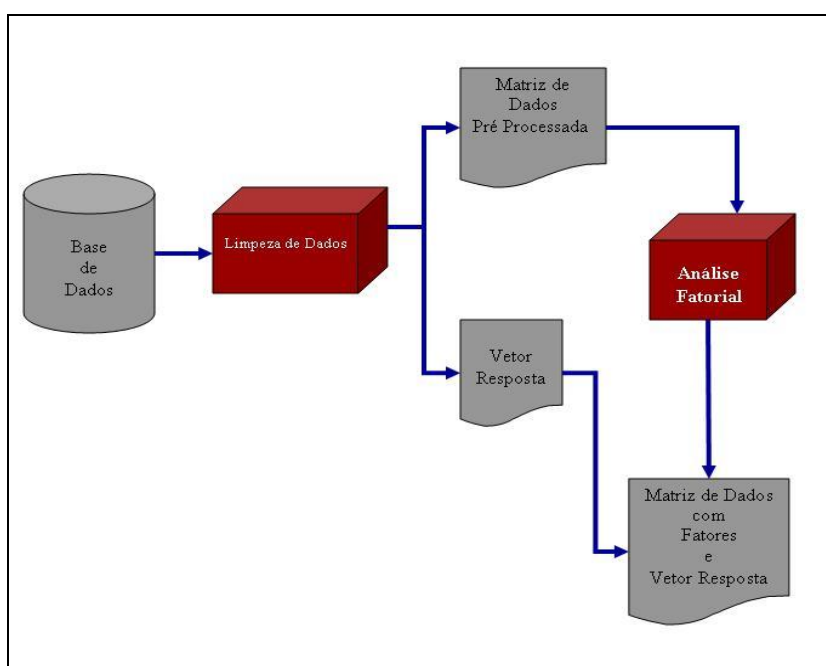


FIGURA 3.11 – PROCESSO DO AGENTE MINERADOR DE DADOS  
 FONTE: O AUTOR

### 3.2.3.2 Agente de Comunicação

Na arquitetura proposta é utilizado um sistema de comunicação confederado, ou seja, os Agentes Inteligentes se comunicam através de um facilitador (Agente de Comunicação). Esse Agente de Comunicação é responsável pela interconexão dos agentes, ele funciona como um centro de comunicação onde todas as mensagens enviadas por qualquer agente passam por ele e são direcionadas para o respectivo destinatário.

A conversação entre os agentes está sempre relacionada com um plano para atingir um objetivo, portanto deve-se especificar qual informação deve ser trocada e com quem, pois um fluxo de informação muito grande pode deteriorar a capacidade de comunicação do meio empregado, o que resultaria numa perda de reatividade do Sistema Multiagentes. Deve lembrar-se que nessas conversações são trocadas informações parciais do processo de tomada de decisão, logo conversações desnecessárias podem prejudicar o processo de tomada de decisões.

A comunicação entre agentes é fundamental para a execução de um Sistema Multiagentes. Ela determina o comportamento em uma sociedade, permitindo que um agente não seja apenas um programa que executa seus comportamentos, mas também um programa que recebe e envia pedidos aos demais agentes.

A troca de mensagens na plataforma JADE realiza-se mediante mensagens FIPA-ACL. JADE disponibiliza um mecanismo assíncrono de mensagens: cada agente possui uma fila de mensagens (caixa de entrada), onde este agente decide o momento de ler estas mensagens. No momento desejado pelo agente este pode ler apenas a primeira mensagem, ou ler as mensagens que satisfazem algum critério.

Para a comunicação entre os agentes se faz necessário a criação de um diagrama de interação entre agentes, o qual explicita a sequência de ações dentro do Sistema Multiagentes, bem como a colaboração entre eles. Na figura 3.12 apresenta o diagrama de interação da arquitetura do Sistema Multiagentes.

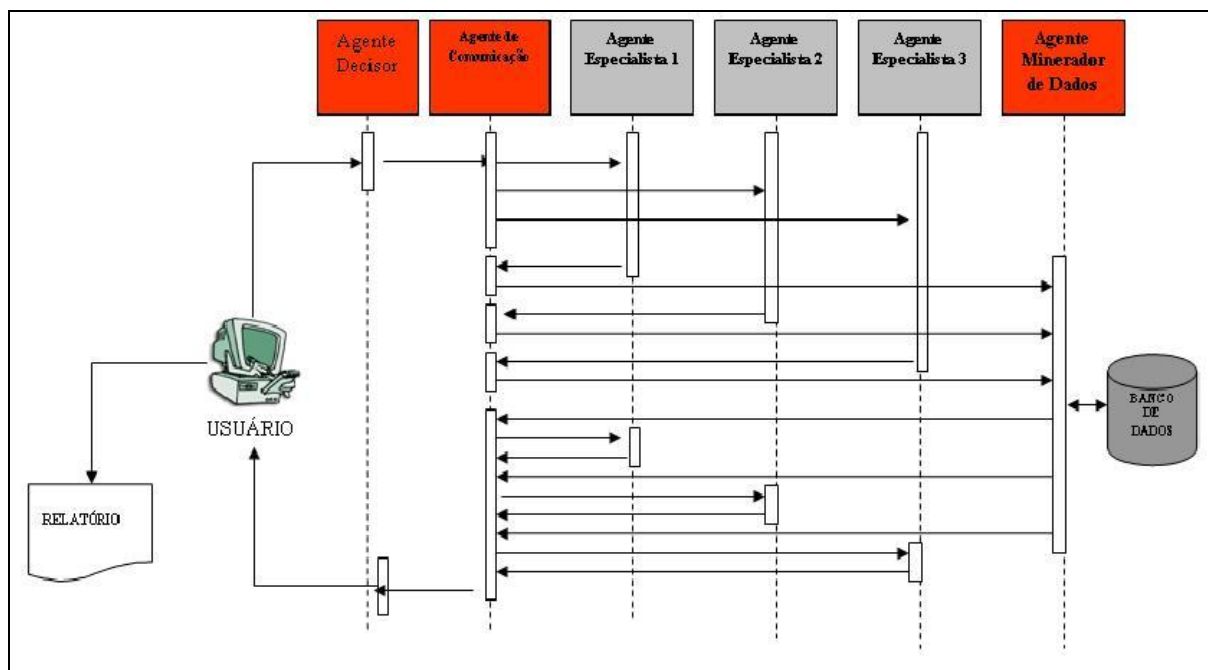


FIGURA 3.12 – DIAGRAMA DE INTERAÇÃO ENTRE OS AGENTES  
 FONTE: O AUTOR

### 3.2.3.3 Agente Decisor

O Agente Decisor é responsável pela interface entre o Usuário com o sistema, sendo assim responsável pelo gerenciamento de todo o sistema. As funções deste agente podem ser divididas em duas funções principais:

#### I. Interface do Sistema com o Usuário

O Agente Decisor faz toda a comunicação do Sistema Multiagentes com o Usuário, desta forma este agente inicia o processo de Reconhecimento de Padrões ou de Reconhecimento de Padrões e Classificação de uma Nova Observação de acordo com a necessidade do usuário.

Quando o processo de Reconhecimento de Padrões é solicitado ao Agente Decisor, o mesmo tem como função iniciar todo o processo Reconhecimento de Padrão, de acordo com a base de dados informada pelo usuário e apresentar ao Usuário os modelos gerados e a Taxa Real de Acerto processada pelos Agentes Especialistas, bem como os fatores gerados pelo Agente Minerados de Dados, bem como suas comunalidades.

Quando o processo de Reconhecimento de Padrões e Classificação de uma Nova Observação é solicitado, o Agente Decisor repete o processo de Reconhecimento de Padrões, porém neste caso o Agente Decisor tem um papel fundamental de ponderar a classificação da nova amostra apresentada pelos Agentes Especialistas com base na Taxa Real de Acerto. Essa ponderação executada pelo Agente Decisor é feita por uma Lógica Nebulosa.

## II. Ponderação de Classificação

Para executar a ponderação de classificação o Agente Decisor utiliza uma Lógica Nebulosa de modelo Mandani, apresentada na figura 3.13

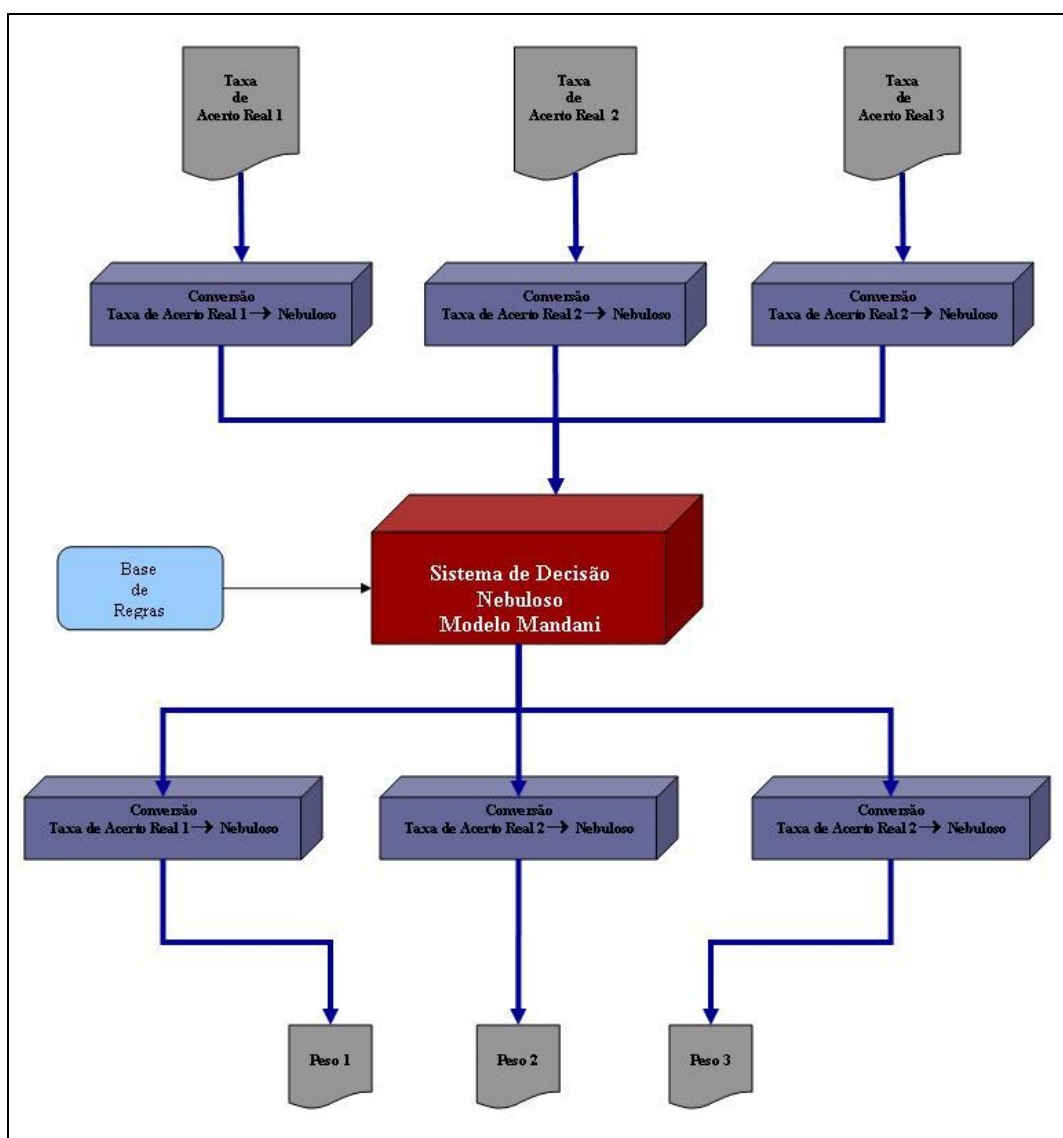


FIGURA 3.13 – PROCESSO DO AGENTE DECISOR  
FONTE: O AUTOR

Para o modelo desenvolvido são utilizadas três variáveis numéricas de entrada contendo os valores das Taxas de Acerto Reais de cada Agente Especialista.

- Taxa Real de Acerto – 1: Regressão Logística
- Taxa Real de Acerto – 2: Escore Quadrático
- Taxa Real de Acerto – 3: Redes Neurais

Foram utilizadas três variáveis de saída contendo os pesos para cada Classificação da Nova Observação obtida pelos Agentes Especialistas.

- Peso – 1 ( $W_1$ ): Regressão Logística
- Peso – 2 ( $W_2$ ): Escore Quadrático
- Peso – 3 ( $W_3$ ): Redes Neurais

As funções de pertinência desenvolvida para as variáveis de entrada, e para as variáveis de saída são apresentadas na figura 3.14 e figura 3.15

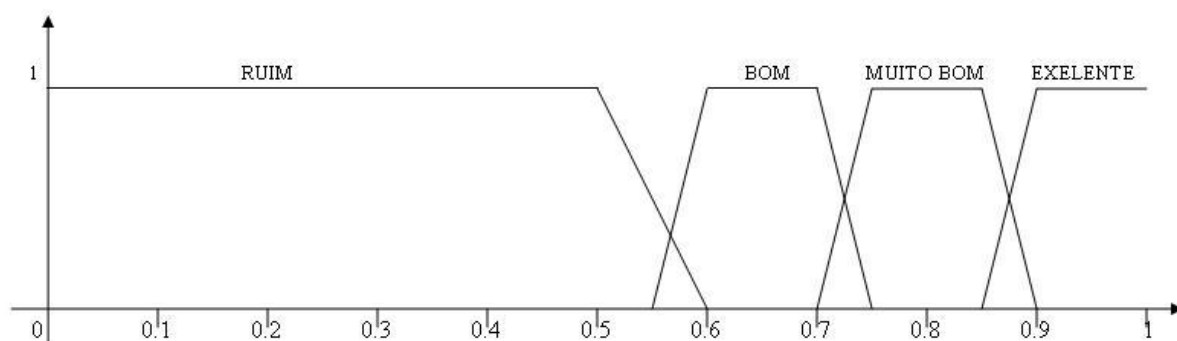


FIGURA 3.14 – FUNÇÃO DE PERTINÊNCIA PARA TAXA REAL DE ACERTO  
 FONTE: O AUTOR

As variáveis de entrada foram divididas em quatro subconjuntos nebulosos considerando uma faixa de valores possíveis no intervalo de [0,1]. Esses subconjuntos foram definidos, conforme descrito a seguir:



- Ruim: escolheu-se uma função trapezoidal com as seguintes parametrizações [0, 0, 0.50, 0.60];
- Bom: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.55, 0.60, 0.70, 0.75];
- Muito Bom: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.70, 0.75, 0.85, 0.90];
- Excelente: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.85, 0.90, 1, 1];

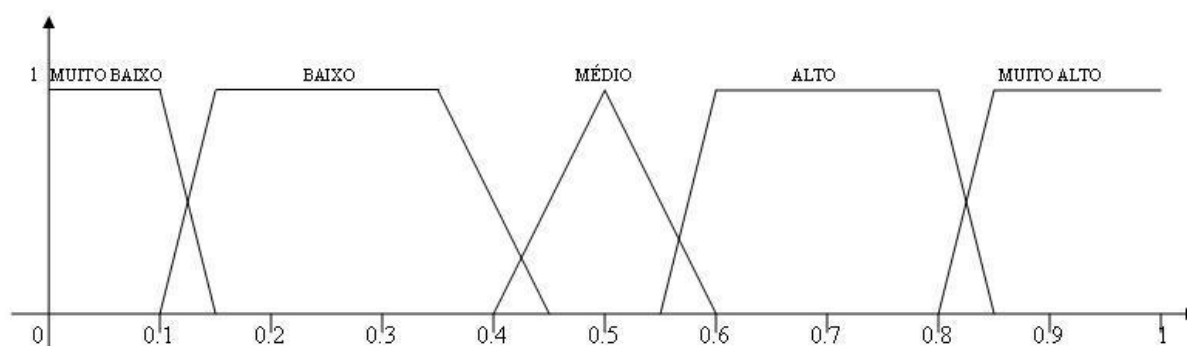


FIGURA 3.15 – FUNÇÃO DE PERTINÊNCIA PARA VARIÁVEL DE SAÍDA  
 FONTE: O AUTOR

As variáveis de saída foram divididas em cinco subconjuntos nebulosos considerando uma faixa de valores possíveis no intervalo de [0,1]. Esses subconjuntos foram definidos, conforme descrito a seguir:

- Taxa de Acerto Muito Baixa: escolheu-se uma função trapezoidal com as seguintes parametrizações [0, 0, 0.10, 0.15];
- Taxa de Acerto Baixa: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.10, 0.15, 0.35, 0.45];
- Taxa de Acerto Médio: escolheu-se uma função triangular com as seguintes parametrizações [0.40, 0.50, 0.60];
- Taxa de Acerto Alto: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.55, 0.60, 0.80, 0.85];
- Taxa de Acerto Muito Alto: escolheu-se uma função trapezoidal com as seguintes parametrizações [0.80, 0.85, 1, 1];

Foram desenvolvidas 64 regras para a base de conhecimento, a seguir é apresentado um exemplo dessas regras:

<Se “Taxa Real de Acerto – 1 ” for “Bom” e Taxa Real de Acerto – 2 ” for “Muito Bom” e Taxa Real de Acerto – 3 ” for “Ruim”> então < “Peso – 1” é “Médio” e “Peso – 2” é “Alto” e “Peso – 3.” é “Muito Baixa”.>

Com base nos pesos apresentados das variáveis de saída e na Classificação da Nova Observação dada por cada Agente Especialista é possível calcular o valor da Resposta Final de Classificação (R.F.C.) dado na equação 3.1.

$$R.F.C = \bar{W}_1.RA_1 + \bar{W}_2.RA_2 + \bar{W}_3.RA_3 \quad (3.1)$$

onde:

- $\bar{W}_i = \frac{W_i}{\sum W_i}$
- $RA_1$  é a classificação da nova observação pelo Agente Especialista 1 – Regressão Logística;
- $RA_2$  é a classificação da nova observação pelo Agente Especialista 2 – Escore Quadrática;
- $RA_3$  é a classificação da nova observação pelo Agente Especialista 1 – Redes Neurais.

E a Regra de Classificação é dada por:

$$SE \begin{cases} R.F.C. > c & \text{Nova Observação} \in \text{Grupo}_1 \\ R.F.C. < c & \text{Nova Observação} \in \text{Grupo}_0 \end{cases} \quad (3.1)$$

onde  $c$  é o valor da faixa de corte do sistema (valor adotado para classificação igual a 0,5).

### 3.2.4 Agentes de Especialistas

Os Agentes Especialistas são detentores do conhecimento específico empregado na análise e solução dos problemas encontrados no domínio de aplicação. Na arquitetura proposta, existem três agentes especialistas responsáveis pelos Métodos de Reconhecimento e Classificação de Padrões.

Os Agentes Especialistas recebem a nova base de dados processada pelo Agente Minerador de Dados, em posse dessa nova base de dados cada Agente Especialista necessita de um pré-processamento de informações a fim de que essa base de dados esteja em um formato necessário para a aplicação das respectivas técnicas de Reconhecimento e Classificação de Padrões. Todos os agentes especialistas utilizam o método de Lachenbruch (*leave-one-out*), para obter as Taxas de Acertos Reais. O fluxograma da figura 3.16 apresenta o processo dos Agentes Especialistas.

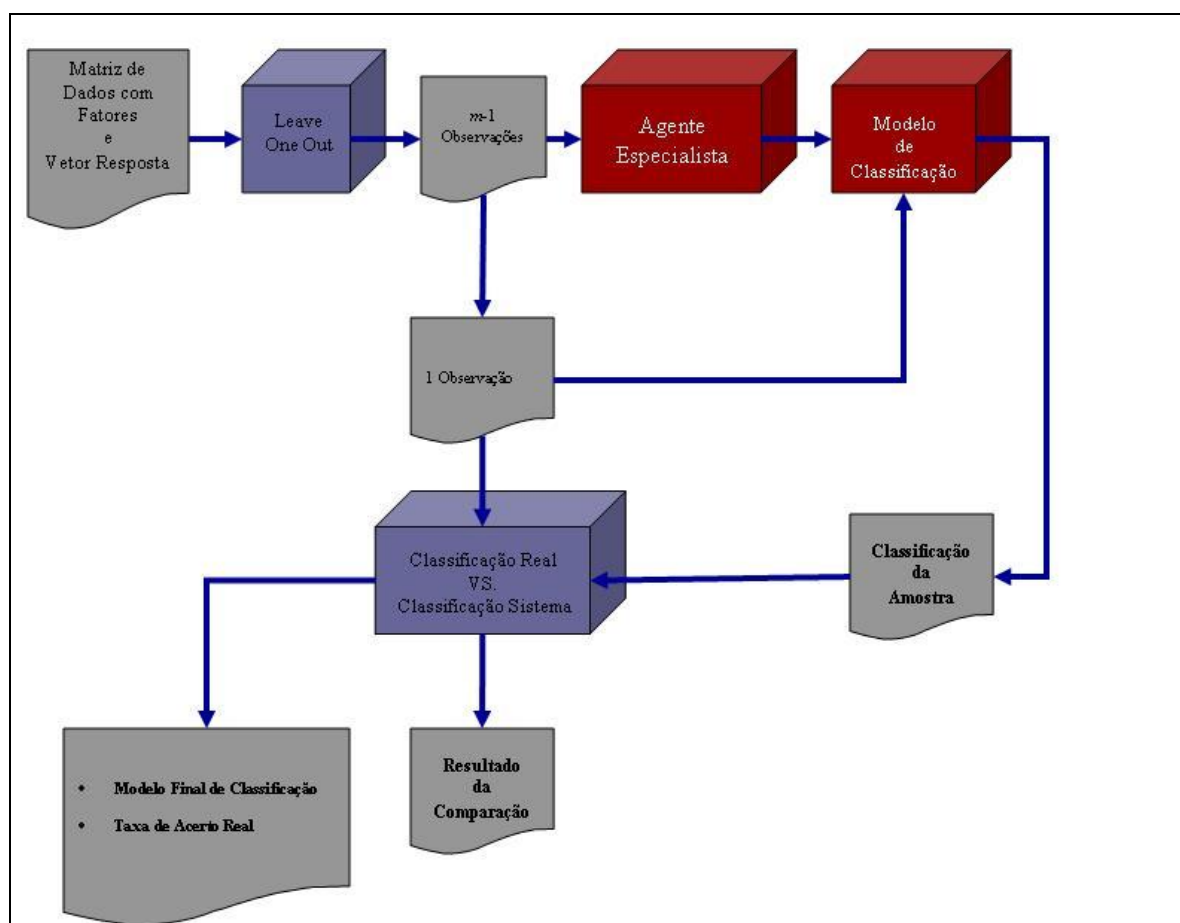


FIGURA 3.16 – PROCESSO DOS AGENTES ESPECIALISTAS  
FONTE: O AUTOR

As técnicas e metodologias de cada Agente Especialista são descritas a seguir.

#### 3.2.4.1 Agente Especialista 1 – Regressão Logística

O Agente Especialista – Regressão Logística, primeiramente aplica uma separação do banco de dados em uma matriz contendo as variáveis independentes do modelo e um vetor resposta que contem as variáveis independentes do modelo. Desta forma estima-se o vetor de parâmetros  $\underline{\beta} = [\beta_0, \beta_1, \dots, \beta_p]$  pelo método da máxima verossimilhança, num contexto de mínimos quadrados não-lineares, aplicando o Algoritmo de Quase-Newton.

Foram gerados  $n$  modelos com ( $n-1$  observações), para teste e validação da técnica. Esse processo é conhecido como método de Lachenbruch (*leave-one-out*). Ao final do processo é ajustado um modelo com as  $n$  observações e que passará a ser considerado como definitivo para o Reconhecimento e Classificação de Padrões.

No modelo de Regressão Logística é necessária a utilização de uma faixa de corte com valores no intervalo  $[0,1]$ , que serve para determinar a classificação de uma nova observação. Foram testadas, neste trabalho, várias faixas de corte contidas no intervalo  $[0,1]$ , porém no sistema adotou-se como valor da faixa de corte a proporção entre os dois grupos.

#### 3.2.4.2 Agente Especialista 2 – Escore Quadrática

O Agente Especialista – Escore Quadrática, necessita primeiramente de uma separação do banco de dados em uma matriz contendo as variáveis independentes e um vetor contendo as variáveis dependentes do modelo. Posteriormente, a matriz das variáveis independentes é dividida em dois conjuntos, cada um deles contendo as variáveis referentes a cada grupo. O vetor com as variáveis dependentes do modelo não são necessárias para a criação das funções de classificação.

A essas matrizes dos grupos é aplicado o método de Lachenbruch (*leave-one-out*) e são criados  $n$  funções de classificações com (com  $n-1$  amostras), dessa forma é obtido a taxa de acerto real dentro de cada grupo e ao final a função de classificação é criada com as  $n$  amostras.

#### 3.2.4.3 Agente Especialista – Redes Neurais

De posse do novo banco de dados o Agente Especialista – Redes Neurais processa as informações do banco de dados.

A implementação da técnica de Redes Neurais foi feita utilizando o algoritmo desenvolvido no software MATLAB e, também, utilizando rotinas oriundas do próprio pacote do MATLAB. O algoritmo aplica o método de Lachenbruch (*leave-one-out*), no treinamento e no teste da rede, que consiste na classificação de uma amostra retirada do banco de dados (com  $n-1$  observações) que originou as regras (LACHENBRUCH, 1975). Esse procedimento é repetido  $n$  vezes.

Para a criação da Rede Neural optou-se por uma rede com treinamento *backpropagation* com correção de erro pelo algoritmo de Levenberg-Marquadt que faz parte do default da função utilizada no MATLAB.

Assim foi necessário fixar como número de neurônios na camada de entrada que é o mesmo que o número de variáveis da base de dados fornecidas pelo Agente Minerador de Dados. O número de variáveis também define o número de camadas ocultas utilizadas, ou seja, número de camadas ocultas +1. Na camada de saída foi fixado um neurônio. Nas camadas ocultas se optou por utilizar uma função de ativação sigmoideal, ou seja, do tipo tangente hiperbólica.

O modelo de Rede Neural implementado sorteia os valores iniciais dos sinápticos e das bias. Assim faz-se necessário apenas fixar os valores das épocas de treinamento e a magnitude do gradiente. Após vários testes optou-se por utilizar 1000 épocas de treinamento e uma magnitude de 0,0001.

Cada modelo de Rede Neural treinado teve um tempo computacional médio de 8 segundos.

## 4 RESULTADOS E DISCUSSÕES

Nos testes da arquitetura do sistema desenvolvido foi utilizado primeiramente um banco de dados teórico contendo oito (8) variáveis dependentes e uma (1) variável independente. Este banco contém 40 observações dessas variáveis sendo que dez (10) fazem parte do grupo 0 e 30 fazem parte do grupo 1. A tabela 4.1 apresenta o banco de dados utilizado para a validação da Arquitetura.

TABELA 4.1 – BANCO DE DADOS PARA TESTE DA ARQUITETURA

Variável 1	Variável 2	Variável 3	Variável 4	Variável 5	Variável 6	Variável 7	Variável 8	Variável Resposta
1,60	1,67	152,17	0,56	0,94	0,69	12,54	26,42	1
8,65	0,13	21,24	4,17	7,02	7,02	13,35	21,11	1
3,17	0,46	84,57	1,00	1,60	1,33	5,59	7,75	1
1,63	1,58	90,15	1,02	1,25	1,06	-0,45	-3,43	1
2,15	0,87	56,5	1,16	1,40	1,5	7,46	7,83	1
1,94	1,06	61,38	0,98	1,50	1,36	0,82	3,49	1
1,97	1,03	84,46	0,40	0,95	1,14	0,45	2,08	1
1,68	1,47	104,91	0,79	1,00	0,97	-0,02	0,24	1
2,89	0,53	59,13	3,86	3,86	1,76	7,75	23,21	1
6,66	0,18	90,02	1,11	1,55	1,68	9,03	6,05	1
1,43	2,32	51,67	0,53	1,06	1,20	0,29	1,42	1
3,60	0,38	91,81	1,59	3,22	1,21	15,53	20,52	1
1,67	1,49	144,64	0,71	1,76	0,70	19,18	84,12	1
1,71	1,40	35,57	1,91	2,13	1,45	3,78	11,15	1
13,44	0,09	95,61	2,39	3,57	2,51	17,49	17,98	1
3,05	0,49	115,49	0,70	0,79	0,68	2,23	2,64	1
21,29	0,05	98,44	1,79	2,07	2,07	21,94	12,84	1
2,24	0,81	87,25	1,11	1,49	1,16	-0,06	0,98	1
3,18	0,46	76,34	2,17	2,18	1,51	10,03	7,30	1
5,71	0,28	125,00	1,29	1,29	1,29	26,06	56,67	1
1,37	2,67	229,32	0,42	0,55	0,52	-0,99	-11,28	1
1,60	1,68	78,97	1,21	1,28	1,12	6,69	29,24	1
3,54	0,39	35,21	2,27	2,65	2,65	15,00	32,73	1
1,47	2,14	87,65	0,60	1,03	1,03	-2,13	-23,46	1
4,37	0,30	84,45	1,59	1,59	1,52	3,79	7,59	1
1,64	1,55	25,26	0,16	1,48	1,48	9,01	77,34	1
1,23	4,31	187,46	0,31	0,36	0,76	0,52	7,67	1
1,25	4,05	86,62	0,36	1,03	1,03	2,07	17,35	1
5,30	0,23	94,38	0,54	1,03	1,24	-3,06	-9,83	1
1,67	1,59	216,77	0,33	0,39	0,29	-1,05	-10,85	0
1,21	4,72	387,18	0,21	0,40	0,39	-23,45	-147,63	0
1,25	4,04	102,84	-0,03	1,00	0,99	2,02	8,06	0
1,08	17,52	63,00	0,24	1,21	1,05	-7,76	-25,24	0
1,17	5,88	219,05	0,36	1,29	0,80	1,62	2,20	0
1,87	1,15	87,61	0,74	1,28	1,10	-40,36	-135,03	0
1,45	2,22	227,74	0,39	0,92	0,42	-50,67	-26,3	0
1,19	5,31	475,96	0,30	0,33	0,27	-0,61	-9,55	0
2,09	0,92	189,03	0,06	0,06	0,03	-0,08	5,61	0
1,31	15,81	154,59	0,75	1,35	1,21	4,72	34,67	0
1,72	1,39	97,87	2,61	2,61	0,61	9,70	32,58	0
1,60	1,67	152,17	0,56	0,94	0,69	12,54	26,42	1

Esse banco de dados reduzido foi utilizado para teste, devido ao tempo computacional de treinamento de cada agente especialista, visto que para validação do Sistema Multiagentes para Reconhecimento e Classificação de Padrões, cada Agente Especialista utiliza o método de Lachenbruch (*leave-one-out*) para criação dos seus modelos.

#### 4.1 – TESTE DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES

Nos testes da Arquitetura proposta, primeiramente foi testada a eficácia do Agente Minerador de Dados. Para testar a eficácia do Agente Minerador de Dados foi comparado a Taxa de Acerto Real de cada método de Reconhecimento e Classificação de Padrões no banco de dados original, com a Taxa de Acerto Real de cada método de Reconhecimento e Classificação de Padrões no banco de dados dos fatores gerados pela Análise Fatorial. As Taxas de Acerto Real com o banco de dados original são apresentadas nos quadros 4.1, 4.2 e 4.3

QUADRO 4.1 – CLASSIFICAÇÃO DOS DADOS ORIGINAIS PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	*7	3
		70,00%	30,00%
	Grupo 1	5	*25
		16,66%	83,33%
<b>Taxa de Acerto Real</b>	<b>80,00%</b>		

\*Percentual de Acerto

QUADRO 4.2 – CLASSIFICAÇÃO DOS DADOS DAS ORIGINAIS PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	*6	4
		60,00%	40,00%
	Grupo 1	2	*28
		6,66%	93,33%
<b>Taxa de Acerto Real</b>	<b>85,00%</b>		

\*Percentual de Acerto

QUADRO 4.3 – CLASSIFICAÇÃO DOS DADOS ORIGINAIS PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	*8 80,00%	2 20,00%
	Grupo 1	6 20,00%	*24 80,00%
<b>Taxa de Acerto Real</b>		<b>80,00%</b>	

\*Percentual de Acerto

Posteriormente foi aplicado o método da Análise Fatorial ao banco de dados obtendo assim os seguintes resultados desta análise.

TABELA 4.2 – ANÁLISE FATORIAL DOS DADOS DE TESTE

Número de Fatores	Autovalores	Proporção da Variância	Variância Cumulativa
1	3,86528	48,316	48,316
2	1,38006	17,251	65,567
<b>3</b>	<b>1,00345</b>	<b>11,815</b>	<b>77,381</b>
4	0,794289	9,929	87,310
5	0,540912	6,761	94,071
6	0,241361	3,017	97,088
7	0,184763	2,310	99,398
8	0,0481656	0,602	100,000

FONTE: O AUTOR

A tabela 4.2, apresenta um resumo das informações da aplicação do Método da Análise Fatorial para os dados de teste. O objetivo da análise é a obtenção de um número de fatores  $m$  com  $m < p$  sendo  $p$  a dimensão do vetor observado. Esses  $m$  fatores são responsáveis por explicar a maior parte da variabilidade das oito (8) variáveis. Neste caso, três fatores foram extraídos, então  $m = 3$ . E para determinar o número de fatores utilizou-se o método de Kaiser, no qual o número de fatores a ser extraído é dado pelo número de autovalores da matriz de correlação que tinham valores maiores que ou iguais a um (1), apresentado na figura 4.1. Juntos, eles respondem por 77,381% da variabilidade dos dados originais.



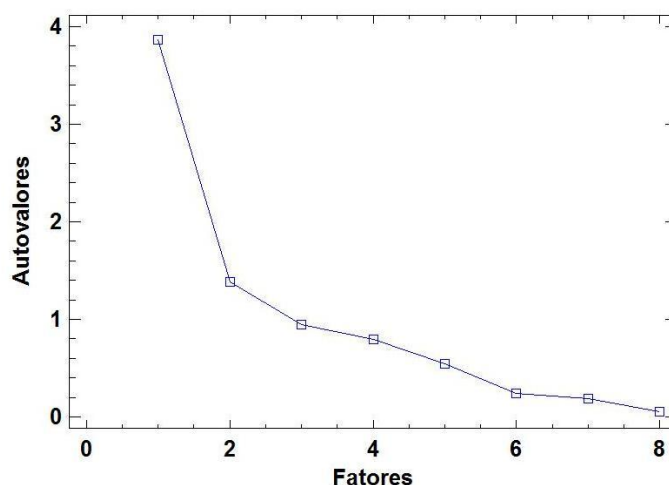


FIGURA 4.1 – SELEÇÃO DO NÚMERO DE FATORES DOS DADOS DE TESTE  
 FONTE O AUTOR

Na tabela 4.3 são apresentados os carregamentos (pesos) dos fatores rotacionados pelo método Varimax, as comunalidades e as porcentagens explicativas.

TABELA 4.3 – CARREGAMENTOS FATORIAIS ROTACIONADA OS DO BANCO DE DADOS DE TESTE

Variáveis	Fator 1	Fator 2	Fator 3	Comunalidades
Variável 1	0,4180	0,1120	<b>0,6049</b>	0,55328
Variável 2	-0,0649	-0,0870	<b>-0,8896</b>	0,80334
Variável 3	<b>-0,5411</b>	-0,4115	-0,0247	0,46283
Variável 4	<b>0,8477</b>	0,1542	0,2848	0,82362
Variável 5	<b>0,9420</b>	0,1515	0,1508	0,93326
Variável 6	<b>0,9150</b>	0,0882	0,1369	0,86384
Variável 7	0,2095	<b>0,8648</b>	0,2301	0,84476
Variável 8	0,0987	<b>0,9463</b>	0,0087	0,90555
<b>Total da Variância</b>				
<b>Explicada (%)</b>	<b>48,31%</b>	<b>17,25%</b>	<b>11,81%</b>	
<b>Total da Variância</b>	<b>48,31%</b>	<b>65,56%</b>	<b>77,38%</b>	
<b>Acumulada (%)</b>				

FONTE: O AUTOR

Em relação aos dados de teste foram identificados apenas três fatores que representam 77,38% da variabilidade, porém esse três fatores estão na direção de maior variabilidade. Os fatores obtidos pela Análise Fatorial são descritos a seguir:

Fator 1: (Variância Explicada 48,31%)

- Variável 3 (Comunalidade = 0,46283 ; Peso = -0,5411)
- Variável 4 (Comunalidade = 0,82362; Peso = -0,8477)
- Variável 5 (Comunalidade = 0,93326; Peso = 0,9420)
- Variável 6 (Comunalidade = 0,86384; Peso = 0,9150)

Fator 2: (Variância Explicada 17,25%)

- Variável 7 (Comunalidade = 0,84476; Peso = 0,8648)
- Variável 8 (Comunalidade = 0,90555; Peso = 0,9463)

Fator 3: (Variância Explicada 11,81%)

- Variável 1 (Comunalidade = 0,55328; Peso = 0,6049)
- Variável 2 (Comunalidade = 0,80334; Peso = -0,8896)

Após a rotação dos fatores foram calculados os escores fatoriais que são utilizados como as variáveis independentes para os métodos de Reconhecimento e Classificação de Padrões. A partir da nova base de dados foram aplicados três métodos de Reconhecimento e Classificação de Padrões, todos eles utilizando o método de Lachenbruch (*leave-one-out*), para obter a Taxa Real de Acerto. Os quadros 4.4, 4.5 e 4.6 apresentam as Classificações e Taxa Real de Acerto dos dados de teste:

QUADRO 4.4 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	8 80,00%	2 20,00%
	Grupo 1	4 13,33%	26 86,66%
<b>Taxa de Acerto Real</b>		<b>85,00%</b>	

\*Percentual de Acerto

QUADRO 4.5 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	8 80,00%	2 20,00%
	Grupo 1	3 10,00%	27 90,00%
<b>Taxa de Acerto Real</b>		<b>87,50%</b>	

\*Percentual de Acerto

QUADRO 4.6 – CLASSIFICAÇÃO DOS DADOS GERADOS PELA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	6 60,00%	4 40,00%
	Grupo 1	0 0,00%	30 100,00%
<b>Taxa de Acerto Real</b>		<b>90,00%</b>	

\*Percentual de Acerto

Em todos os casos com a aplicação da Análise Fatorial apresentou uma melhora nos métodos de Reconhecimento e Classificação de Padrões, a comparação destes resultados é apresentada no quadro 4.7

QUADRO 4.7 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE TESTE

		Taxa Real de Acerto		
		Regressão Logística	Escore Quadrático	Redes Neurais
Banco de Dados	Sem Agente de Mineração	80,00%	85,00%	80,00%
	<b>Com Agente de Mineração</b>	<b>85,00%</b>	<b>87,50%</b>	<b>90,00%</b>

Após a aplicação do Agente Minerador testou-se a eficácia do Agente Decisor no gerenciamento da síntese dos resultados dos Agentes Especialistas, desta forma foi criado uma Lógica Nebulosa a fim de ponderar a classificação feita pelos Agentes Especialistas com suas respectivas Taxa de Acerto Real. Assim para testar a eficácia do todo o sistema foi Sistema Multiagentes para Reconhecimento e

Classificação de Padrões, primeiramente foi aplicado a um banco de dados teóricos, posteriormente foi utilizado o método de Lachenbruch (*leave-one-out*) em todo o sistema como visto na figura 4.2.

Para utilizar o método de Lachenbruch (*leave-one-out*), foi utilizado a seguinte sequência:

1. Na base de dados contendo  $n$  observações e omite-se uma observação;
2. Com  $(n-1)$  observações aplica-se a Análise Fatorial, com o intuito de gerar os  $p$  fatores que serão usados pelos Agentes Especialistas;
3. Com  $p$  fatores os Agentes Especialistas geram os modelos de Reconhecimento e Classificação de Padrões e a Taxa Real de Acerto;
4. Aplica-se os modelos gerados pelos Agentes Especialistas a observação que foi omitida;
5. Com base nas classificações da observação omitida e na Taxa Real de Acerto, o Agente Decisor pondera a classificação da nova observação, obtendo assim uma classificação final;
6. A classificação final é comparada com a classificação real da observação omitida;
7. Repete-se os passos 1 até 6, até que todas as observações tenham sido classificadas;
8. Seja  $n_1$  o número de observações do grupo 1, e  $n_2$  o número de observações do grupo 0, então a Taxa Real de Acerto do Sistema será dado por:

$$\text{Taxa Real de Acerto} = \frac{\text{Acerto do Sistema}}{n_1 + n_2}$$

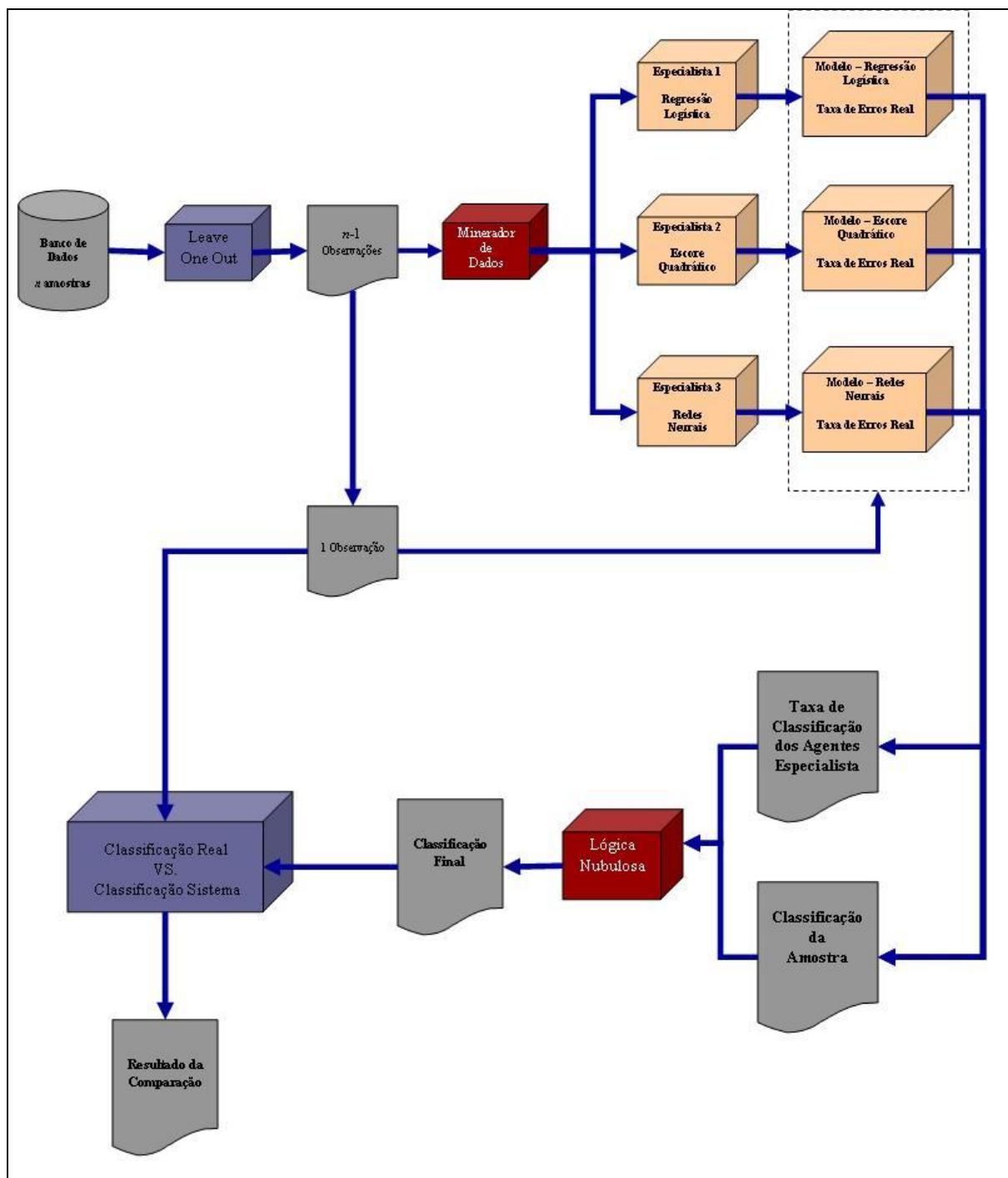


FIGURA 4.2 – MÉTODO DE VALIDAÇÃO DA ARQUITETURA MULTIAGENTES  
 FONTE: O AUTOR

O processo de teste da eficácia da arquitetura do Sistema Multiagentes de Reconhecimento e Classificação de Padrões é apresentada no quadro 4.8.

QUADRO 4.8 – CLASSIFICAÇÃO DOS DADOS DE TESTE PELO AGENTE DECISOR COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Grupo 0	Grupo 1
Classificação Real	Grupo 0	8 80,00%	2 40,00%
	Grupo 1	0 0,00%	30 100,00%
<b>Taxa de Acerto Real</b>		<b>95,00%</b>	

Sistema de pós-processamento nebuloso melhorou o desempenho global do Sistema Multiagentes, tendo uma Taxa Real de Acerto de 95,00%, a comparação dos métodos de Reconhecimento e Classificação de Padrões é apresentada no quadro 4.9.

QUADRO 4.9 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE TESTE

		Taxa Real de Acerto		
		Regressão Logística	Escore Quadrático	Redes Neurais
Banco de Dados	Sem Agente de Mineração	80,00%	85,00%	80,00%
	Com Agente de Mineração	85,00%	87,50%	90,00%
	<b>Agente Decisor</b>	<b>95,00%</b>		

Com base no quadro 4.9, observa-se que com a inserção da técnica de Análise Fatorial como pré-processamento de dados, obteve-se um melhor desempenho dos métodos de Reconhecimento e Classificação de Padrões.

O mesmo pode-se observar com a inserção da decisão nebulosa que incrementou o desempenho final das técnicas de Reconhecimento e Classificação de Padrões isoladas. Desta forma após os testes do Sistema Multiagentes de Reconhecimento e Classificação de Padrões sem executados, optou-se por validar a eficácia do sistema em duas bases de dados reais de problemas de Engenharia de Potência.

## 4.2 VALIDAÇÃO DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES

A arquitetura do Sistema Multiagentes de Reconhecimento e Classificação de Padrões foi aplicada na identificação e classificação de transformadores da rede elétrica de acordo com o nível de distorção harmônica total de tensão e na identificação das quedas de estruturas metálicas, e comparado com resultados dos trabalhos de Souza (2008), Martins (2010) e Wazen (2011), com o intuito de validar a eficácia do sistema.

### 4.2.1 Validação do Agente Minerador de Dados

O Agente Minerador de Dados permite a automação do pré-processamento de dados sem a necessidade da expertise do especialista. Esse pré-processamento de dados tem como objetivo:

- Reduzir o volume de dados;
- Reduzir o tempo computacional dos métodos de Reconhecimento e Classificação de Padrões;
- Agrupar as variáveis originais que têm uma determinada associação e estão nas direções de maior variabilidade;

Todos esses objetivos citados devem ser alcançados sem reduzir significativamente no desempenho dos métodos de Reconhecimento e Classificação de Padrões.

Desta forma a medida para avaliar a eficácia do Agente Minerador de Dados foi a Taxa Real de Acerto dos métodos de Reconhecimento e Classificação de Padrões antes e depois da aplicação da Análise Fatorial.

O Agente Minerador de Dados foi aplicado aos dois conjuntos de dados para validação do Sistema Multiagentes, os resultados serão apresentados dentro de cada conjunto de dados.

#### 4.2.1.1 Caso I – Distorções Harmônicas de Tensão

Primeiramente foram aplicados os métodos de Reconhecimento e Classificação de Padrões ao conjunto de dados sem nenhum tratamento, todos os métodos aplicados utilizam o método de Lachenbruch (*leave-one-out*), para obter a Taxa Real de Acerto, desta forma os resultados obtidos são apresentados nos Quadros 4.10, 4.11 e 4.12.

QUADRO 4.10 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	237 *84,34%	44 15,65%
	Problemático	56 71,79%	22 *28,20%
<b>Taxa de Acerto Real</b>		<b>72,14%</b>	

\*Percentual de Acerto

QUADRO 4.11 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	180 *64,05%	101 35,94%
	Problemático	26 33,33%	52 *66,67%
<b>Taxa de Erro Real</b>		<b>64,63%</b>	

\*Percentual de Acerto

QUADRO 4.12 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	273 *97,15%	8 2,85%
	Problemático	53 67,94%	25 *32,05%
<b>Taxa de Erro Real</b>		<b>83,01%</b>	

\*Percentual de Acerto



Posteriormente foi aplicado o método da Análise Fatorial ao banco de dados referente às variáveis independentes das distorções harmônicas de tensão obtendo assim os seguintes resultados desta análise.

TABELA 4.4 – ANÁLISE FATORIAL DOS DADOS DAS DISTORÇÕES HARMONICAS DE TENSÃO

Número de Fatores	Autovalores	Proporção da Variância	Variância Cumulativa
1	4,74934	27,93%	27,93%
2	2,30975	13,58%	41,52%
3	1,58098	9,30%	50,82%
4	1,31581	7,74%	58,56%
5	1,13126	6,65%	65,21%
<b>6</b>	<b>1,02187</b>	<b>6,01%</b>	<b>71,22%</b>
7	0,926229	5,44%	76,67%
8	0,807517	4,75%	81,42%
9	0,779209	4,58%	86,01%
10	0,602181	3,54%	89,55%
11	0,551315	3,24%	92,79%
12	0,445736	2,62%	95,41%
13	0,305962	1,80%	97,21%
14	0,215793	1,26%	98,48%
15	0,122075	0,71%	99,20%
16	0,090702	0,53%	99,74%
17	0,044276	0,26%	100,00%

FONTE: O AUTOR

A tabela 4.4, apresenta um resumo das informações da aplicação do Método da Análise Fatorial para os dados das Distorções Harmônicas de Tensão. O objetivo da análise é a obtenção de um número de fatores  $m$  com  $m < p$ , sendo  $p$  a dimensão do vetor observado. Esses  $m$  fatores são responsáveis por explicar a maior parte da variabilidade das 17 variáveis. Neste caso, seis (6) fatores foram extraídos, para determinar o número de fatores utilizou-se o método de Kaiser, no qual o número de fatores a ser extraído é dado pelo número de autovalores da matriz de correlação que tinham valores maiores que ou iguais a 1, apresentado na figura 4.3. Juntos, eles respondem por 71,22% da variabilidade dos dados originais.

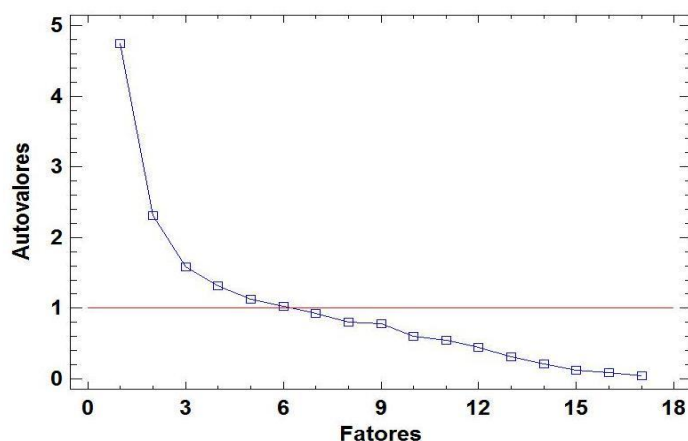


FIGURA 4.3 – SELEÇÃO DO NÚMERO DE FATORES DOS DADOS DAS DHTV  
 FONTE O AUTOR

Na tabela 4.5 são apresentados os carregamentos (pesos) dos fatores rotacionados pelo método Varimax, as comunalidades e as porcentagens explicativas.

TABELA 4.5 – CARREGAMENTOS FATORIAIS ROTACIONADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO

Variáveis	Fator 1	Fator 2	Fator 3	Fator 4	Fator 5	Fator 6
Superintendência Regional	0,0419	-0,2127	<b>0,5643</b>	0,0395	-0,1511	0,3304
Tensão Primária	-0,0495	-0,0662	<b>0,8395</b>	0,0407	0,0171	-0,0276
Fases	0,1720	0,2522	<b>0,6763</b>	-0,3494	0,0923	0,0449
Potência	0,1854	0,0088	<b>-0,7072</b>	0,3451	0,0123	0,1382
Número de Consumidores	<b>0,8237</b>	0,4923	-0,0591	0,1626	0,0157	0,0282
Residencial 0/50 kWh	<b>0,8079</b>	-0,0813	-0,0651	-0,0131	0,0075	-0,0097
Residencial 50/100 kWh	<b>0,9466</b>	0,0243	-0,0250	-0,0177	-0,0590	-0,0021
Residencial 100/150 kWh	<b>0,8788</b>	0,2855	-0,0751	-0,0161	-0,0779	-0,0260
Residencial 150/200 kWh	<b>0,6289</b>	0,6568	-0,0934	0,0281	-0,0462	0,0310
Residencial 200/300 kWh	0,2597	<b>0,9010</b>	-0,1007	0,0663	0,0643	-0,0085
Residencial 300/500 kWh	-0,0061	<b>0,8771</b>	-0,0972	0,0611	0,1752	0,0232
Residencial 500/1000 kWh	-0,0902	0,3028	0,0031	0,1522	<b>0,7652</b>	-0,0262
Residencial >1000 kWh	-0,0081	-0,0390	-0,0841	-0,0316	<b>0,8785</b>	-0,0330
Comércio e Poder Público	-0,0658	0,0159	0,0726	<b>0,7465</b>	0,0195	0,1867
Industrial	-0,2306	0,0063	0,1911	-0,2471	-0,1524	-0,0163
Rural	-0,0591	0,0376	-0,0019	0,1836	-0,0456	<b>0,8092</b>
Iluminação Pública	-0,1113	-0,0246	-0,0051	<b>0,5846</b>	-0,0337	-0,5323
<b>Total da Variância Explicada (%)</b>	<b>27,93</b>	<b>13,58</b>	<b>9,30</b>	<b>7,74</b>	<b>6,65</b>	<b>6,01</b>
<b>Total da Variância Acumulada (%)</b>	<b>27,93</b>	<b>41,52</b>	<b>50,82</b>	<b>58,56</b>	<b>65,21</b>	<b>71,22</b>

FONTE: O AUTOR

Em relação aos dados referentes as Distorções Harmônicas de Tensão foram identificados seis fatores que representam 71,22% da variabilidade, porém esse seis fatores estão na direção de maior variabilidade, os fatores obtidos pela Análise Fatorial são descritos a seguir:

Fator 1: Consumo Baixo (Variância Explicada 27,93%)

- Número de Consumidores (Comunalidade = 0,9520 ; Peso = 0,8237)
- Residencial 0/50 kWh (Comunalidade = 0,6639; Peso = 0,8079)
- Residencial 50/100 kWh (Comunalidade = 0,9011; Peso = 0,9466)
- Residencial 100/150 kWh (Comunalidade = 0,8666; Peso = 0,8788)
- Residencial 150/200 kWh (Comunalidade = 0,8396; Peso = 0,6289)
- Residencial 200/300 kWh (Comunalidade = 0,8991; Peso = 0,2597)

Fator 2: Fator Consumo Médio (Variância Explicada 13,58%)

- Residencial 200/300 kWh (Comunalidade = 0,8138; Peso = 0,9010)
- Residencial 300/500 kWh (Comunalidade = 70,92; Peso = 0,8771)

Fator 3: Fator Operacional (Variância Explicada 9,30%)

- Superintendência Regional (Comunalidade = 0,5123; Peso = 0,5643)
- Tensão Primária (Comunalidade = 0,7144; Peso = 0,8395)
- Fases(Comunalidade = 0,6833, Peso = 0,6763)
- Potencia (Comunalidade = 0,6729, Peso = -0,7072)

Fator 4: Fator Comércio/Poder Público/Iluminação Pública( Variância Explicada 7,74%)

- Comércio e Poder Público (Comunalidade = 0,6025; Peso = 0,0726)
- Iluminação Pública (Comunalidade = 0,6393; Peso = 0,5846)

Fator 5: Fator Consumo Alto(Variância Explicada 6,65%)

- Residencial 500/1000 kWh (Comunalidade = 0,7825; Peso = 0,7652)
- Residencial >1000 kWh (Comunalidade = 0,6025; Peso = 0,8785)

Fator 6: Fator Rural(Variância Explicada 6,01%)

- Rural(Comunalidade = 0,6956; Peso = 0,8092)

A variável Industrial apresenta uma comunalidade de 0,1743, isso indica que essa variável pode ser descartada na análise visto que não tem variância explicativa, significativa para o modelo. Lembrando-se que as medições deste banco de dados foram realizadas no secundário de transformadores de distribuição, focando-se na rede de baixa tensão, onde não se verificam cargas significativas do tipo Industrial.

Com base nos fatores obtidos após a rotação dos fatores foram calculado os escores fatoriais e estes são utilizados como as variáveis independentes para os métodos de Reconhecimento e Classificação de Padrões. A partir da nova base de dados foram aplicados três métodos de Reconhecimento e Classificação de Padrões, todos eles utilizando o método de Lachenbruch (*leave-one-out*), para obter a Taxa Real de Acerto. Os quadros com as Classificações e Taxa Real de Acerto são apresentadas a seguir:

QUADRO 4.13 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	*241 85,76%	40 14,23%
	Problemático	53 67,94%	25 *32,05%
<b>Taxa de Acerto Real</b>		<b>74,09%</b>	

\*Percentual de Acerto

QUADRO 4.14 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	*177 62,98%	104 37,02%
	Problemático	13 16,67%	65 *83,33%
<b>Taxa de Acerto Real</b>		<b>67,41%</b>	

\*Percentual de Acerto

QUADRO 4.15 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	*273	8
		97,15%	2,85%
	Problemático	53	25
		67,94%	*32,05%
<b>Taxa de Acerto Real</b>	<b>83,01%</b>		

\*Percentual de Acerto

Em todos os casos com a aplicação da Análise Fatorial apresentou uma melhora no desempenho dos métodos de Reconhecimento e Classificação de Padrões, a comparação destes resultados é apresentada no quadro 4.16

QUADRO 4.16 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO

		Taxa Real de Acerto		
		Regressão Logística	Escore Quadrático	Redes Neurais
Distorções Harmônicas de Tensão	Sem Agente de Mineração	72,15%	64,63%	83,01%
	<b>Com Agente de Mineração</b>	<b>74,09%</b>	<b>67,41%</b>	<b>83,01%</b>

Com base no quadro 4.17, pode-se observar que a aplicação da Análise Fatorial como pré-processamento da informação teve uma melhora de 1,94% no método de Regressão Logística, 2,78% no método de Escore Quadrático, e não teve melhora no método de Redes Neurais.

#### 4.2.1.2 Caso II – Quedas de Estruturas Metálicas de Linhas de Transição

Primeiramente foram aplicados os métodos de Reconhecimento e Classificação de Padrões ao conjunto de dados das Quedas de Estruturas sem nenhum tratamento, todos os métodos aplicados utilizam o método de Lachenbruch (*leave-one-out*), para obter a Taxa Real de Acerto, desta forma os resultados obtidos são apresentados nos quadros a 4.17, 4.18 e 4.19.

QUADRO 4.17 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*23 *52,27%	21 47,73%
	Não Queda	16 11,67%	*121 *88,32%
<b>Taxa de Acerto Real</b>		<b>79,56%</b>	

\*Percentual de Acerto

QUADRO 4.18 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*33 *75,00%	11 25,00%
	Não Queda	13 9,49%	*124 *90,51%
<b>Taxa de Erro Real</b>		<b>86,74%</b>	

\*Percentual de Acerto

QUADRO 4.19 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS PELO MÉTODO DE REDES NEURAIIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*23 *52,27%	21 47,72%
	Não Queda	12 13,20%	*125 *86,80%
<b>Taxa de Erro Real</b>		<b>81,76%</b>	

\*Percentual de Acerto

Posteriormente foi aplicado o método da Análise Fatorial obtendo assim os seguintes resultados:

TABELA 4.6 – ANÁLISE FATORIAL DOS DADOS DAS QUEDAS DE ESTRUTURAS

Número de Fatores	Autovalores	Proporção da Variância	Variância Cumulativa
1	1,99686	19,96%	19,96%
2	1,64037	16,40%	36,37%
3	1,44306	14,43%	50,80%
<b>4</b>	<b>1,10751</b>	<b>11,07%</b>	<b>61,87%</b>
5	0,907899	9,07%	70,95%
6	0,829768	8,29%	79,25%
7	0,708422	7,08%	86,33%
8	0,609269	6,09%	92,43%
9	0,384674	3,84%	96,27%
10	0,372167	3,72%	100,00%

FONTES: O AUTOR

A tabela 4.6, apresenta um resumo das informações da aplicação do Método da Análise Fatorial para os dados de Quedas de Estruturas. O objetivo da análise é a obtenção de um número de fatores  $m$  com  $m < p$ , sendo  $p$  a dimensão do vetor observado. Esses  $m$  fatores que são responsáveis por explicar a maior parte da variabilidade das 10 variáveis. Neste caso, quatro fatores foram extraídos, para determinar o número de fatores utilizou-se o método de Kaiser, no qual o número de fatores a ser extraído é dado pelo número de autovalores da matriz de correlação que tinham valores maiores que ou iguais a 1, apresentado na figura 4.4. Juntos, eles respondem por 61,878% da variabilidade dos dados originais.

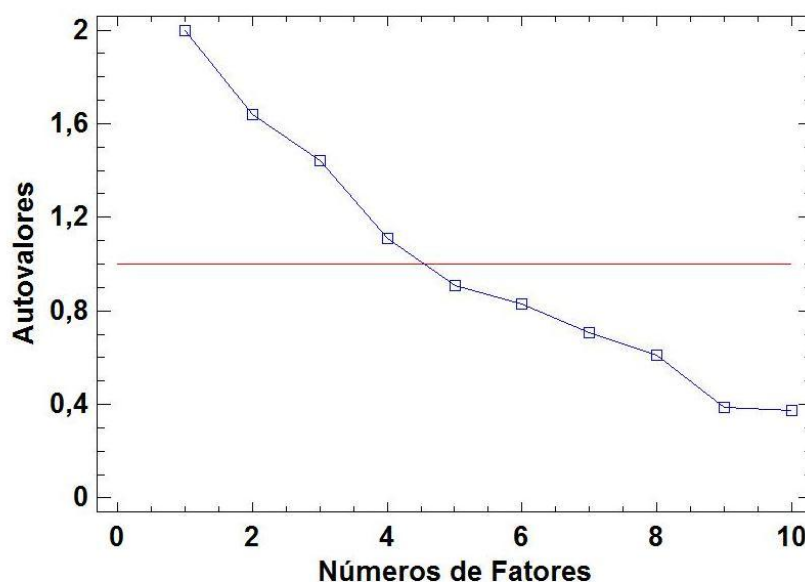


FIGURA 4.4 – SELEÇÃO DO NÚMERO DE FATORES DAS QUEDAS DE ESTRUTURAS  
 FONTE O AUTOR

Na tabela 4.7 são apresentados os carregamentos (pesos) dos fatores rotacionados pelo método Varimax, as comunalidades e as porcentagens explicativas.

TABELA 4.7 – CARREGAMENTOS FATORIAIS ROTACIONADOS DAS QUEDAS DE ESTRUTURAS

Variáveis	Fator 1	Fator 2	Fator 3	Fator 4	Comunalidades
Tensão	0,1010	<b>0,8056</b>	-0,0351	-0,1184	0,6745
Vento	-0,0056	<b>0,6884</b>	-0,3440	0,1907	0,6287
Relevo	0,2159	-0,0419	<b>0,7402</b>	-0,0403	0,5979
Massa de Ar Fria	<b>0,7385</b>	0,0863	0,3223	0,1324	0,6743
Massa de Ar Quente	<b>-0,8353</b>	0,0187	0,0158	0,1082	0,7096
Temperatura	<b>0,6598</b>	-0,2628	-0,3092	-0,0555	0,6032
Terreno	-0,1927	-0,0689	<b>0,7221</b>	-0,0667	0,5677
Capacidade	-0,1958	-0,1844	-0,0243	<b>0,7573</b>	0,6465
Função	0,1329	0,1106	-0,0758	<b>0,7572</b>	0,6091
Fundação	-0,2834	<b>0,6091</b>	0,1385	-0,0729	0,4759
<b>Total da Variância Explicada (%)</b>	<b>19,96%</b>	<b>16,40%</b>	<b>14,43%</b>	<b>11,07%</b>	
<b>Total da Variância Acumulada (%)</b>	<b>19,96%</b>	<b>36,37%</b>	<b>50,80%</b>	<b>61,87%</b>	

FONTE: O AUTOR

Em relação aos dados referentes as Quedas de Estruturas foram identificados apenas quatro fatores que representam 61,87% da variabilidade, porém esse quatro fatores estão na direção de maior variabilidade, os fatores obtidos pela Análise Fatorial são descritos a seguir:

Fator 1: Fator Climático (Variância Explicada 19,96%)

- Massa de Ar Fria (Comunalidade = 0,6745; Peso = 0,7385)
- Massa de Ar Quente (Comunalidade = 0,7096; Peso = -0,8353)
- Temperatura (Comunalidade = 0,6032, Peso = 0,6598)

Fator 2: Fator Determinação da Estrutura (Variância Explicada 16,40%)

- Tensão (Comunalidade = 0,6745 ; Peso = 0,8056)
- Vento (Comunalidade = 0,6287; Peso = 0,6884)
- Fundação (Comunalidade = 0,4759; Peso = 0,6091)



Fator 3: Fator Geográfico (Variância Explicada 14,43%)

- Revelo (Comunalidade = 0,5979; Peso = 0,7402)
- Terreno (Comunalidade = 0,5677; Peso = 0,7221)

Fator 4: Fator Características da LT (Variância Explicada 11,07%)

- Capacidade (Comunalidade = 0,6465; Peso = 0,7573)
- Função (Comunalidade = 0,6091; Peso = 0,7572)

Com base nos fatores obtidos após a rotação dos fatores foi calculado os escores fatoriais e estes são utilizados como as variáveis independentes para os métodos de Reconhecimento e Classificação de Padrões. A partir da nova base de dados foram aplicados três métodos de Reconhecimento e Classificação de Padrões, todos eles utilizando o método de Lachenbruch (*leave-one-out*), para obter a Taxa Real de Acerto. Os quadros com as Classificações e Taxa Real de Acerto são apresentadas nos quadros 4.20, 4.21 e 4.22.

QUADRO 4.20 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DA REGRESSÃO LOGÍSTICA COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*36 81,81%	8 18,19%
	Não Queda	28 21,89%	*109 78,11%
<b>Taxa Real de Acerto</b>		<b>80,11%</b>	

\*Percentual de Acerto

QUADRO 4.21 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DO ESCORE QUADRÁTICO COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*39 88,63%	5 11,37%
	Não Queda	17 14,60%	*120 85,40%
<b>Taxa Real de Acerto</b>		<b>87,84%</b>	

\*Percentual de Acerto

QUADRO 4.22 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	*23 52,27%	21 47,72%
	Não Queda	12 8,75%	*125 91,24%
<b>Taxa Real de Acerto</b>		<b>81,76%</b>	

\*Percentual de Acerto

Em todos os casos com a aplicação da Análise Fatorial apresentou uma melhora no desempenho dos métodos de Reconhecimento e Classificação de Padrões, a comparação destes resultados é apresentada no quadro 4.23.

QUADRO 4.23 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS QUEDAS DE ESTRUTURAS

		Taxa Real de Acerto		
		Sem Agente de Mineração	Com Agente de Mineração	Com Agente de Mineração
Queda de Estruturas	Sem Agente de Mineração	79,56%	86,74%	81,76%
	<b>Com Agente de Mineração</b>	<b>80,11%</b>	<b>87,84%</b>	<b>81,76%</b>

Com base no quadro 4.23, pode-se observar que a aplicação da Análise Fatorial como pré-processamento da informação teve uma melhora de 0,55% no método de Regressão Logística, 1,10% no método de Escore Quadrático, e não teve melhora no método de Redes Neurais. No quadro, 4.24 é apresentado um resumo da aplicação da técnica de Análise Fatorial nos dados de validação.

QUADRO 4.24 – RESUMO DA APLICAÇÃO DA ANÁLISE FATORIAL COMO AGENTE MINERADOR DE DADOS

		Taxa Real de Acerto		
		Sem Agente de Mineração	Com Agente de Mineração	Com Agente de Mineração
Queda de Estruturas	Sem Agente de Mineração	79,56%	86,74%	81,76%
	<b>Com Agente de Mineração</b>	<b>80,11%</b>	<b>87,84%</b>	<b>81,76%</b>
Distorções Harmônicas de Tensão	Sem Agente de Mineração	72,15%	64,63%	83,01%
	<b>Com Agente de Mineração</b>	<b>74,09%</b>	<b>67,41%</b>	<b>83,01%</b>

Com base no quadro 4.24, pode-se observar que em todos os casos a aplicação da Análise Fatorial representou uma melhora nos métodos de Reconhecimento e Classificação de Padrões. No banco de dados referente às Quedas de Estruturas, foi reduzido o número de variáveis de entrada de dez para quatro. No banco de dados referente às Distorções Harmônicas de Tensão, houve uma redução de 17 variáveis para seis fatores. Em ambos os casos houve uma no desempenho dos métodos de Reconhecimento e Classificação de Padrões. Desta forma o Agente Minerador de Dados mostrou-se eficaz visto que houve:

- Redução do volume de dados
- Agrupamento das variáveis originais em fatores
- Melhor desempenho computacional devido à redução do volume de dados;
- Melhor performance dos métodos de Reconhecimento e Classificação de Padrões
- Uma análise da correlação entre as variáveis e sua importância na variabilidade dos conjuntos de dados.

#### 4.2.1 Validação do Decisor

O Agente Decisor é responsável pela comunicação do Sistema Multiagentes com o Usuário, porém tem um papel fundamental no gerenciamento da síntese dos resultados apresentados pelos Agentes Especialistas. Para a validação do Agente Decisor foi utilizado o método de Lachenbruch (*leave-one-out*) em todo o Sistema Multiagentes. Os resultados para cada conjunto de dados serão apresentados a seguir.

##### 4.2.1.1 – Caso I: Distorções Harmônicas de Tensão

Após a aplicação do Agente Minerador utilizou-se os resultados dos Agentes Especialistas, para a criação de uma Lógica Nebulosa a fim de ponderar a classificação feita pelos Agentes Especialistas com suas respectivas Taxa de Acerto Real. No banco de dados das Distorções Harmônicas de Tensão foi utilizado o método de Lachenbruch (*leave-one-out*) em todo o sistema e comparado com o

valor da classificação real dos dados, a classificação e a Taxa Real de Acerto do Agente Decisor é apresentado no quadro 4.25.

QUADRO 4.25 – CLASSIFICAÇÃO DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO APÓS A APLICAÇÃO DA LÓGIA NEBULOSA

		Classificação Prevista	
		Bom	Problemático
Classificação Real	Bom	*277	4
		98,57%	1,42%
	Problemático	22	*52
		32,46%	67,53%
<b>Taxa Real de Acerto</b>		<b>91,64%</b>	

Com base no quadro 4.25, o Agente Decisor apresentou uma Taxa Real de Acerto de 91,64%, essa Taxa Real de Acerto foi comparada com os resultados das técnicas de Reconhecimento e Classificação de Padrões isoladamente, com e sem a aplicação da Análise Fatorial. Esta comparação é apresentada no quadro 4.26

QUADRO 4.26 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DE DISTORÇÕES HARMÔNICAS DE TENSÃO

Banco de Dados		Taxa Real de Acerto		
		Regressão Logística	Escore Quadrático	Redes Neurais
Distorções Harmônicas de Tensão	Sem Agente de Mineração	79,56%	86,74%	81,76%
	Com Agente de Mineração	80,11%	87,84%	81,76%
	<b>Agente Decisor</b>	<b>91,64%</b>		

No quadro 4.26, pode-se observar que a aplicação do Agente Decisor no gerenciamento da síntese dos modelos de Reconhecimento e Classificação de Padrões obteve-se uma melhora média de 8,40% em relação as técnicas Isoladas.

#### 4.2.1.2 - Caso II - Quedas de Estruturas Metálicas de Linhas de Transição

No banco de dados das Quedas de Estruturas foi utilizado o método de Lachenbruch (*leave-one-out*) em todo o sistema utilizando os resultados dos Agentes Especialistas e suas respectivas Taxas Real de Acerto para a criação de uma Lógica Nebulosa e comparado com o valor da classificação real dos dados, a classificação e a Taxa Real de Acerto do Agente Decisor é apresentado no quadro 4.27.

QUADRO 4.27 – CLASSIFICAÇÃO DOS DADOS DAS QUEDAS DE ESTRUTURAS APÓS A APLICAÇÃO DA ANÁLISE FATORIAL PELO MÉTODO DE REDES NEURAIS COM TAXA DE ACERTO REAL

		Classificação Prevista	
		Queda	Não Queda
Classificação Real	Queda	37	7
		84,09%	15,90%
	Não Queda	10	*127
		7,29%	92,70%
<b>Taxa Real de Acerto</b>		<b>90,60%</b>	

\*Percentual de Acerto

Com base no quadro 4.27, o Agente Decisor apresentou uma Taxa Real de Acerto de 90,60%, essa Taxa Real de Acerto foi comparada com os resultados das técnicas de Reconhecimento e Classificação de Padrões isoladamente, com e sem a aplicação da Análise Fatorial. Esta comparação é apresentada no quadro 4.28.

QUADRO 4.28 – COMPARAÇÃO DOS MÉTODOS DE RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES DOS DADOS DAS QUEDAS DE ESTRUTURAS

Banco de Dados		Taxa Real de Acerto		
		Regressão Logística	Escore Quadrático	Redes Neurais
Distorções Harmônicas de Tensão	Sem Agente de Mineração	72,15%	64,63%	83,01%
	Com Agente de Mineração	74,09%	67,41%	83,01%
	<b>Agente Decisor</b>	<b>90,60%</b>		

No quadro 4.28, pode-se observar que a aplicação do Agente Decisor no gerenciamento da síntese dos modelos de Reconhecimento e Classificação de Padrões obteve-se uma melhora média de 15,76% em relação as técnicas Isoladas. Um resumo da aplicação da Lógica Nebulosa no pós-processamento da informação é apresentado no quadro 4.29.

QUADRO 4.29 – RESUMO DA APLICAÇÃO DA LÓGICA NEBULOSA COMO AGENTE DECISOR

		Taxa Real de Acerto		
Queda de Estruturas	Sem Agente de Mineração	79,56%	86,74%	81,76%
	Com Agente de Mineração	80,11%	87,84%	81,76%
	<b>Agente Decisor</b>	<b>90,60%</b>		
Distorções Harmônicas de Tensão	Sem Agente de Mineração	72,15%	64,63%	83,01%
	Com Agente de Mineração	74,09%	67,41%	83,01%
	<b>Agente Decisor</b>	<b>91,64%</b>		

Com base no quadro 4.29 pode-se observar que o Agente Decisor consegue sintetizar e integrar os desempenhos das três Técnicas de Reconhecimento e Classificação de Padrões testadas, melhorando assim o desempenho global do sistema.

Dessa forma, o sistema de decisão nebuloso desenhado consegue inferir uma classificação final aproveitando as vantagens individuais de cada técnica para suprir as deficiências pontuais de alguma classificação errada obtida por uma determinada técnica.

#### 4.3 – RESULTADOS DA VALIDAÇÃO DO SISTEMA MULTIAGENTES PARA RECONHECIMENTO E CLASSIFICAÇÃO DE PADRÕES

A arquitetura proposta foi aplicada em dois bancos de dados reais de problemas na Engenharia de Potência. Souza (2008) e Martins (2010) que utilizam

separadamente técnicas de Reconhecimento e Classificação de Padrões no banco de dados das Distorções Harmônicas de Tensão, porém nos dois trabalhos os bancos de dados foram separados previamente em cinco superintendências regionais. A comparação entre a classificação apresentada nos dois trabalhos e do Sistema Multiagentes para Reconhecimento e Classificação de Padrões é apresentado no quadro 4.30.

QUADRO 4.30 – COMPARAÇÃO DOS RESULTADOS DOS DADOS DAS DISTORÇÕES HARMÔNICAS DE TENSÃO

Regional	<i>Rough sets</i>		<i>Regressão Logística</i>	<i>Escore Quadrático</i>	Sistema Multiagentes
	Indefinições	Taxa Real de Acerto	Taxa Real de Acerto	Taxa Real de Acerto	Taxa Real de Acerto
SDC	9,4%	37,90%	56,62%	56,25%	<b>91,64%</b>
SDO	27,4%	62,20%	72,58%	53,96%	
SDN	44,4%	65,70%	61,90%	64,51%	
SDT	24,3%	71,40%	70,27%	67,56%	
SDL	19,5%	79,60%	85,15%	75,00%	

Com base no quadro 4.30, tem-se que o Sistema Multiagentes para Reconhecimento e Classificação de Padrões apresentou um resultado superior as técnicas aplicadas separadamente. Vale ressaltar que o Sistema Multiagentes efetuou o Reconhecimento e Classificação de Padrões em toda a base de dados, descartando assim expertise do especialista.

No banco de dados de Quedas de Estruturas o resultado do Sistema Multiagentes para Reconhecimento e Classificação de Padrões foi comparado aos resultados apresentando por Wazen (2011), que utilizou duas técnicas de Reconhecimento e Classificação de Padrões. Os resultados das comparações são apresentados no quadro 4.31.

QUADRO 4.31 – COMPARAÇÃO DOS RESULTADOS DOS DADOS DAS QUEDAS DE ESTRUTURAS

	<i>Rough sets</i>		<i>Regressão Logística</i>	Sistema Multiagentes
	Indefinições	Taxa Real de Acerto	Taxa Real de Acerto	Taxa Real de Acerto
Quedas de Estruturas	76,24%	69,76%	80,11%	<b>90,60%</b>

Neste banco de dados o Sistema Multiagentes obteve um resultado melhor do que das técnicas aplicadas separadamente por Wazen (2011), o que comprova a eficácia do Sistema Multiagentes no processo de Mineração de Dados e no Gerenciamento dos métodos de Reconhecimento e Classificação de Padrões.



## 5 CONCLUSÕES

Neste trabalho buscou-se apresentar uma visão geral sobre a importância dos métodos de Reconhecimento e Classificação de Padrões em problemas reais. Desta forma destacamos a necessidade de métodos cada vez mais eficientes para suprir problemas reais de Reconhecimento e Classificação de Padrões. Visto que em muitos casos a aplicação de apenas um método de Reconhecimento e Classificação de Padrões não supre essa necessidade, então um sistema que integre vários métodos se torna cada vez mais importante.

Com o intuito de utilizar a arquitetura de Sistema Multiagentes para integrar e gerenciar os métodos de Reconhecimento e Classificação de Padrões foi feito um estudo sobre as definições e a fundamentação teórica dos Sistemas Multiagentes, bem como as principais vantagens. Foram também investigados os principais desafios enfrentados pela comunidade científica na concepção e execução dos Sistemas Multiagentes. Podemos destacar que os principais desafios enfrentados estão relacionados a escolha de design específicos para a criação dos agentes, problemas relacionados a interoperabilidade e a ontologia, bem como a escolha de design para a concepção de agentes autônomos. Desta forma foi apresentado o atual estado da arte sobre as aplicações dos Sistemas Multiagentes na Engenharia de Potência.

A fim de validar a arquitetura proposta foram utilizados dois bancos de dados contemplando problemas reais de Reconhecimento e Classificação de Padrões na Engenharia de Potência.

Para uma análise do banco de dados referente as Distorções Harmônicas de Tensão foi apresentado uma visão geral da questão da qualidade de energia elétrica bem como seus principais distúrbios, sua definições, características e as principais normas técnicas sobre limites para os distúrbios da Qualidade de Energia Elétrica. Para o banco de dados foram utilizadas informações históricas de medições de harmônicos de tensão, realizadas nos transformadores do sistema de distribuição da COPEL. Adicionalmente, foram consideradas na elaboração do modelo, algumas características técnicas e operacionais dos transformadores monitorados. Estas informações associaram-se aos dados de consumo, tipo e quantidade de

consumidores, tensão de operação e potência nominal do transformador, tipo de ligação, dentre outras informações julgadas importantes.

Para a análise do banco de dados das Quedas de Estruturas foram descritas as informações gerais sobre linhas de transmissão, considerando as partes que as compõem bem como as principais características destes componentes. Estas informações foram necessárias para identificação da relação das características com o problema, levantando e apresentando quais dados foram relevantes para os estudos relacionados aos esforços adicionais às quais as estruturas estão sujeitas.

No processo de construção da Arquitetura de um Sistema Multiagentes para Reconhecimento e Classificação de Padrões foi necessário definir a função de cada agente inteligente afim de que o Sistema Multiagentes integrasse não somente métodos de Reconhecimento e Classificação de Padrões, mas que gerenciasse desde o pré-processamento do banco de dados até ponderação da classificação de cada método de Reconhecimento e Classificação de Padrões, desta forma foram definidos três agentes fundamentais na arquitetura: Agente Minerador de Dados, Agentes Especialistas e Agente Decisor.

Num primeiro momento foi utilizado um Agente Minerador de dados responsável pela filtragem dos dados e aplicação da Análise Fatorial como uma técnica de pré-processamento de informação com o intuito de que o Sistema Multiagentes tivesse a possibilidade de acessar o banco de dados de forma autônoma, realizando assim uma limpeza de dados, redução do volume de dados, porém sem interferir negativamente na performance dos métodos de Reconhecimento e Classificação de Padrões. Com base nos escores fatoriais rotacionados pode-se observar que a Análise Fatorial agrupou as variáveis com mais correlação em fatores com direções de maior variabilidade e também descartando as variáveis não são significativas para a variabilidade do modelo.

Os resultados com a aplicação da Análise Fatorial no pré-processamento dos dados foi comparado com a aplicação das técnicas de Reconhecimento e Classificação de Padrões no banco de dados originais, desta forma podendo comprovar a eficácia da do Agente Minerador de Dados nos dois bancos de dados em estudo.

Os métodos de Reconhecimento e Classificação de Padrões foram testados e avaliados antes de serem implementados como Agentes Especialistas. Esses métodos foram selecionados por terem apresentados melhores resultados nos

bancos de dados de teste e também por já terem sido aplicados nos mesmos bancos de dados utilizados para validação.

Como a Arquitetura proposta tem como intuito não apenas integrar os métodos de Reconhecimento e Classificação de Padrões, mas que gerenciasse a síntese dos resultados dos Agentes Especialistas. Assim o Agente Decisor teve como principal objetivo o pós-processamento dos resultados apresentados pelos Agentes Especialistas.

O Agente Decisor utiliza uma Lógica Nebulosa do modelo Mandani para ponderar as Taxas Reais de Acerto com a classificação de uma nova observação, obtendo assim um resultado final que contempla as principais vantagens de cada método de Reconhecimento e Classificação de Padrões, suprimindo assim as deficiências pontuais de cada método.

A aplicação da Lógica Nebulosa foi comparada com os métodos de Reconhecimento e Classificação de Padrões aplicados separadamente nos bancos de dados podendo assim comprovar sua eficácia.

De forma individual cada Agente Inteligente apresentou um resultado satisfatório, visto assim que o Sistema Multiagentes integrando o Agente Minerador de Dados os Agentes Especialistas e o Agente Decisor agrega tanto as vantagens do pré-processamento, do Reconhecimento e Classificação de Padrões e do pós-processamento dos resultados, tendo assim um resultado global melhor do que o resultado de cada Agente Inteligente agindo separadamente.

Por fim as principais contribuições científicas deste trabalho são:

- Utilização da Análise Fatorial como pré-processamento de dados para os métodos de Reconhecimento e Classificação de Padrões;
- Criação de uma Lógica Nebulosa no pós-processamento dos resultados dos métodos de Reconhecimento e Classificação de Padrões;
- Estruturação da Mineração de dados em um Sistema Multiagentes para Reconhecimento e Classificação de Padrões.

Fica então como sugestões para trabalhos futuros:

- Testar a Arquitetura do Sistema Multiagentes de Reconhecimento e Classificação de Padrões em outros problemas de Reconhecimento e Classificação de Padrões.
- Utilizar outros métodos de Reconhecimento e Classificação de Padrões como Agentes Especialistas.
- Combinar outras regras dentro do Agente Decisor.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA – ANEEL. **Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional**. PRODIST, módulo 8, ago. 2007.

AOKI, A.R. **Uma Proposta de Arquitetura Multiagentes para Operação de Sistemas Elétricos**. Itajubá, 2003. Tese de Doutorado - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do de Itajubá.

AOKI, A.R. **Notas de Aula da Disciplina Inteligência Artificial**, Universidade Federal do Paraná, Curitiba, PR, Curso Inteligência Artificial Aplicada, 2005.

BAKER, P. C.; CATTERSON, V. M.; MCARTHUR, S. D. J.; **Integrating an Agent-based Wireless Sensor Network within an Existing Multi-Agent Condition Monitoring System**. 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

CATTERSON, V. M.; RUDD, S. E.; MCARTHUR, S. D. J.; MOSS, G.; **On-line Transformer Condition Monitoring through Diagnostics and Anomaly Detection**. 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

CHAVES NETO, A.; **Análise Multivariada Aplicada a Pesquisa: Notas de aula**. Departamento de Estatística, Universidade Federal do Paraná, Curitiba, 1997.

CHIGUTI. **Aplicação da Análise Multivariada na Caracterização dos Municípios Paranaenses Segundo suas Produções Agrícolas**. Dissertação (Mestrado) — UFPR, 2005.

COPEL – Companhia Paranaense de Energia. Disponível em <http://www.copel.com/pagcopel.nsf>

COURY, D. V.; THORP, J. S.; HOPKINSON, K. M.; BIRMAN, K. P.; **An agent-based current differential relay for use with a utility intranet**. IEEE Transaction on Power Delivery, vol. 17, no. 1, pp. 47–53, Jan. 2002.

DAVIDSON, E. M.; MCARTHUR, S. D. J.; MCDONALD, J. R.; CUMMING, T.; WATT, I.; **Applying multi-agent system technology in practice: Automated management and analysis of SCADA and digital fault recorder data**. IEEE Transaction on Power System. vol. 21, no. 2, pp. 559–567, May 2006.

DAVIDSON, E. M.; MCARTHUR, S. D. J.; **Exploiting Multi-agent System Technology within an Autonomous Regional Active Network Management System**. 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

DIMEAS, A. L.; HATZIARGYRIOU, N. D.; **Operation of a multi-agent system for microgrid control.** IEEE Transaction on Power System, vol. 20, no. 3, pp. 1447–1455, Aug. 2005.

DIMEAS, A. L.; HATZIARGYRIOU, N. D.; **Agent based control of Virtual Power Plants.** 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

DU, W.; CHEN, Z.; WANG, H. F.; DUNN, R.; **Feasibility of online collaborative voltage stability control of power systems.** IET Gener. Transm. Distrib., 2009, Vol. 3, Iss. 2, pp. 216–224.

EDWARDS, D.; SRIVASTAVA, S. K.; CARTES, D. A.; SIMMONS, S.; WILDE, N.; **Implementation and Validation of a Multi-Level Security Model Architecture.** 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

ESQUARCINI, R. **Classificação e Ranqueamento dos Municípios Paranaenses Segundo suas Políticas Setoriais Através da Análise Multivariada.** Dissertação (Mestrado) — UFPR, 2005.

FERBER, J. **Multi-Agent Systems: An Introduction to Distributed Artificial Intelligence.** Addison-Wesley, 1999. 509 p.

FAUSETT, L. V. **Fundamentals of Neural Networks - Architectures, Algorithms, and Applications.** Prentice Hall International Editions. USA, New Jersey, 1994.

FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. **From data mining to knowledge discovery: an overview.** Advances in Knowledge Discovery & Data Mining, 1996 pp. 1-34.

Foundation for Intelligent Physical Agents (FIPA). Disponível em <http://www.fipa.org/>.

FRANKLIN, S.; GRAESSER, A.; **Is it an agent or just a program?**, Proc. 3rd Int. Workshop Agent Theories, Architectures, and Languages 1996, New York: Springer-Verlag.

FUCHS, R. D. **Transmissão de Energia Elétrica: Linhas Aéreas – Volume 1.** Rio de Janeiro: Livros Técnicos e Científicos Editora, 1977.

FUKUI, C.; KUDO, H.; KODA, J.; YABE, K.; TOMITA, Y.; **A cooperative protection system with an agent model.** IEEE Transaction on Power Delivery, vol. 13, no. 4, pp. 1060–1066, Oct. 1998.

FURTADO, M. E. et al. **Ranqueamento de faxinais do estado do paraná através da análise fatorial.** Revista Ciências Exatas e Naturais, v. 5, p. 85–99, 2003.

GASSER, L.; BOND, A. H. **Readings in Distributed Artificial Intelligence.** San Mateo: Morgan Kaufmann, 1988.

GEHRKE, O.; BINDNER, H.; **Building a test platform for agents in power system control: Experience from SYSLAB.** 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

GIOVANINI, R.; HOPKINSON, K.; COURY, D. V.; THORP, J. S.; **A primary and backup cooperative protection system based on wide area agents.** IEEE Transaction on Power Delivery, vol. 21, no. 3, pp. 1222–1230, Jul. 2006.

HÄGER, U.; LEHNHOFF, S.; REHTANZ, C.; WEDDE, H. F.; **Multi-Agent System for Coordinated Control of FACTS devices.** 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

HAIR, J. F. et al. **Applied Multivariate Statistical Analysis.** New York: Prentice Hall, 2009.

HAYKIN, S. **Redes Neurais: Princípios e Prática.** 2ª Edição, Porto Alegre: Bookman, 2002.

HILLS, M. **Applied Statistics.** [S.l.]: Book Review, 1977.

HOPKINSON, K.; WANG, X.; GIOVANINI, R.; THORP, J.; BIRMAN, K.; COURY, D.; **EPOCHS: A platform for agent-based electric power and communication simulation built from commercial off-the-shelf components,** IEEE Transaction on Power System, vol. 21, no. 2, pp. 548–558, May 2006.

HUANG, K.; SRIVASTAVA, S. K.; CARTES, D. A.; SLODERBECK, M.; **Intelligent Agents Applied to Reconfiguration of Mesh Structured Power Systems.** 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

**IEC Communications Networks and Systems in Substations.** 2005, Document IEC 61850.

**IEC Energy Management System Application Program Interface (EMS-API) - Part 301:** Common Information Model (CIM) Base, 2005, document IEC 61970-301.

Java Agent Development Framework (JADE). Disponível em <http://jade.cse.it/>.

JOHNSON, R. A.; WICHERN, D. W.; **Applied Multivariate Statistical Analysis.** 4. ed. New Jersey: Prentice-Hall, inc., 1998.

KACHIGAN, S. K. **Multivariate statistical analysis.** New York: Radius Press, 1991.

KAISER, H. F. **The varimax criterion for analytic rotation in factor analysis.** Psychometrika, v. 23, p. 187–200, 1958.

KORBIK, A.; MCARTHUR, S. D. J.; AULT, G. W.; BURT, G. M.; MC-DONALD, J. R.; **Enabling active distribution networks through decentralized autonomous network management.** in Proc. 18th Int. Conf. Electricity Distribution (CIRED), Turin, June 2005.

KORITAROV, V. S.; **Real-world market representation with agents.** IEEE Power and Energy Magazine, vol. 2, no. 4, pp. 39–46, Jul. 2004.

LACHENBRUCH, P. A. **Discriminant Analysis.** New York, Halfner: 1975.

LEE, C.C. **Fuzzy Logic in Control Systems: Fuzzy Logic Controller – Part I.** IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, 1990a. p. 404-418.

LEE, C.C. **Fuzzy Logic in Control Systems: Fuzzy Logic Controller – Part II.** IEEE Transactions on Systems, Man, and Cybernetics, Vol. 20, No. 2, 1990b. p. 419-435.

LIM, I-H.; LEE, S-J.; CHOI, M-S.; CROSSLEY, P.; **Multi-Agent System-based Protection Coordination of Distribution Feeders.** 14th International Conference on Intelligent System Applications to Power Systems, ISAP 2007 November 4 - 8, 2007, Kaohsiung, Taiwan.

LIM, I-H.; CHOI, M-S.; LEE, S-J.; KIM, T. W.; **Intelligent Distributed Restoration by Multi-Agent System Concept in DAS.** 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

LIMA, J. D. **A análise Econômico-financeira de Empresas Sob a Ótica da Estatística Multivariada.** Dissertação (Mestrado em Métodos Numéricos em Engenharia) – Setores de Tecnologia e de Ciências Exatas, Universidade Federal do Paraná, Curitiba, 2002.

LO, Y. L.; WANG, C. H.; LU, C. N.; **A Multi-agent Based Service Restoration in Distribution Network with Distributed Generations.** 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

MAES, P.; **Artificial life meets entertainment: Life-like autonomous agents.** Commun. ACM, vol. 38, no. 11, pp. 108–114, 1995.

MAMDANI, E.H. **Application of Fuzzy Logic to Approximate Reasoning Using Linguistic Synthesis.** IEEE Transactions on Computers, Vol. c-26, No. 12, 1977. p. 1182-1191.

MARQUES, A. F. **Aplicação da análise multivariada na infraestrutura e no desempenhadas escolas públicas do ensino fundamental e médio pertencentes ao núcleo regionalde educação de paranavaí.** Acta Scientiarum, v. 32, p. 75–81, 2010.



MARTINS, R.; **Classificação de Transformadores de Distribuição de Energia Elétrica quanto à DHTV Usando Rough Sets**. Curitiba, 2009. Dissertação de Mestrado - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Paraná.

MCARTHUR, S. D. J.; MCDONALD, J. R.; HOSSACK, J. A.; **A multi-agent approach to power system disturbance diagnosis**. Autonomous Systems and Intelligent Agents in Power System Control and Operation (Power Systems), C. Rehtanz, Ed. New York: Springer-Verlag, Jul. 2003, pp. 75–99.

MCARTHUR, S. D. J.; DAVIDSON, E. M.; DUDGEON, G. J. W.; MC-DONALD, J. R.; **Toward a model integration methodology for advanced applications in power engineering**, IEEE Transaction on Power System, vol. 18, no. 3, pp. 1205–1206, Aug. 2003.

MCARTHUR, S, D, J,; STRACHAN, S, M,; JAHN, G,; **The design of a multiagent transformer condition monitoring system**. IEEE Transaction on Power System. vol. 19, no. 4, pp. 1845–1852, Nov. 2004.

MCARTHUR, S.D.J.; DAVIDSON, E. M.; CATTERSON, V. M.; DIMEAS, A. L.; HATZIARGYRIOU, N. D.; PONCI, F.; FUNABASHI, T.; **Multi-Agent Systems for Power Engineering Applications—Part I: Concepts, Approaches, and Technical Challenges**: IEEE Transactions on Power Systems, vol. 22, n. 4, nov. 2007.

MINGOTI, S. A. **Análise de Dados Através de Métodos de Estatística Multivariada: Uma Abordagem Aplicada**. Belo Horizonte: Editora UFMG, 2005.

MOMOH, J. A.; ALFRED, K.; XIA, Y.; **Framework for Multi-Agent System (MAS) Detection and Control of Arcing of Shipboard Electric Power Systems**. 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

MULLER, S. I. M. G. **Comparação entre os métodos de máxima verossimilhança, distância mínima e o método de Fisher para o reconhecimento de padrões em imagens coloridas**. Curitiba, 1997. 116 f. Dissertação (Mestrado em Geodésica) – Setor de Ciências Geodésicas, Universidade Federal do Paraná.

MÜLLER, S. I. M. G.; CHAVES NETO, A. **Programa de técnicas integradas para análise e avaliação de fornecedores**. RNTI - Revista Negócios e Tecnologia da Informação, v. 2, p. 95–104, 2007.

OPERADOR NACIONAL DO SISTEMA ELÉTRICO – ONS. **Padrões de Desempenho da Rede Básica**, submódulo 2.2, ago. 2002.

OZERENKO, A.; GYUL'MALIEV, A.; GAGARIN, S. **Factorial analysis in selecting the zeolite catalyst for 2,6-dimethylnaphthalene synthesis**. Coke and Chemistry, v. 50, p. 200–206, 2007.

PANDYA, A.S. & MACY, R. B. **Pattern Recognition with Neural Networks in C++**. CRC Press, Inc. Florida Atlantic University, USA, Boca Raton, Florida, 1996.

PARK, S.J.; LIM, J.T.; **Modelling and control of agent-based power protection system using supervisors**. Proc. Inst. Elect. Eng., Control Theory Appl., vol. 153, no. 1, pp. 92–98, Jan. 2006.

QUEIROZ, J. et al. **Geochemical characterization of heavy metal contaminated area using multivariate factorial kriging**. Environmental Geology, v. 55, p. 95–105, 008.

RENCHER, A. C. **Methods of Multivariate Analysis**. New York: Wiley-Interscience, 2002.

ROSA, M.A.; MIRANDA, V.; MATOS, M.; SHEBLÉ, G.; LEITE DA SILVA, A.M.; **Intelligent Agent-based Environment to Coordinate Maintenance Schedule Discussions**. 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

RUSSELL, S.; NORVIG, P.; **Artificial Intelligence: A Modern Approach**. Englewood Cliffs, NJ: Prentice-Hall, 1995.

SHENG, S.; LI, K. K.; CHAN, W. L.; XIANGJUN, Z.; XIANZHONG, D.; **Agent-based self-healing protection system**. IEEE Transaction on Power Delivery, vol. 21, no. 2, pp. 610–618, Apr. 2006.

SILBERSCHATZ, A.; TUZHILIN, A.: **On subjective measures of interestingness in knowledge discovery**. Proceedings of the First International Conference on Knowledge Discovery and Data Mining 1, 1995, p. 275-281.

SOUZA, W. E.; **Análise das Distorções Harmônicas de Tensão a partir de Características dos Transformadores e de Dados de Consumo**. Curitiba, 2008. Dissertação de Mestrado - Programa de Pós-Graduação em Métodos Numéricos em Engenharia, Universidade Federal do Paraná.

SPEARMAN, C. **General intelligence objectively determined and measured**. American Journal of Psychology, v. 15, p. 201–293, 1904.

VAN SICKEL, J. H.; LEE, K. Y. **Distributed Discrete Event and Psuedo Real-time Combined Simulation for Multi-agent Controlled Power Plants**. 15th International Conference on Intelligent System Applications to Power Systems, ISAP 2009 November 8 - 12, 2009, Curitiba, Brazil

VLACHOGIANNIS, J. G.; HATZIARGYIOU, N. D.; LEE, K. Y.; **Ant colony system-based algorithm for constrained load flow problem**. IEEE Transaction on Power System, vol. 20, no. 3, pp. 1241–1249, Aug. 2005.

WAZEN, R. N. **Proposta de Aplicação da Manutenção Baseada em Confiabilidade para Linhas de Transmissão da COPEL**. Monografia de Especialização da Universidade Tecnológica Federal do Paraná, 2008.

WAZEN, R. N. **Avaliação da Suscetibilidade de falhas em Estruturas Metálicas de Linhas de Transmissão**, Curitiba, 2011. Dissertação de Mestrado - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal do Paraná.

WIDERGREN, S. E.; ROOP, J. M.; GUTTROMSON, R. T.; HUANG, Z.; **Simulating the dynamic coupling of market and physical system operations**. in Proceedings of the IEEE Power Engineering Society General Meeting, Jun. 2004, pp. 748–753.

WOOLDRIDGE, M.; JENNINGS, N. R.; **Pitfalls of agent oriented development**. in Proc. 2nd International Conference Autonomous Agents, 1998, pp. 385–391.

WOOLDRIDGE, M.; WEISS, G.; **Intelligent Agents**. Multi-agent Systems, Cambridge, MA: The MIT Press, Apr. 1999, pp. 3–51.

ZADEH, L.A. Fuzzy Sets. **Information and Control**, Vol. 8, 1965. p. 338-353.

ZADEH, L.A. **Outline of a New Approach to the Analysis of Complex Systems and Decision Processes**. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-3, No. 1, 1973. p. 28-44.