

LUCIMAR MARIA FOSSATTI DE CARVALHO

**MODELOS DE APRENDIZAGEM
PARA SISTEMAS NEURO-DIFUSOS**

**FLORIANÓPOLIS
2007**

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA
ELÉTRICA

MODELOS DE APRENDIZAGEM
PARA SISTEMAS NEURO-DIFUSOS

Tese submetida à
Universidade Federal de Santa Catarina
como parte dos requisitos para
obtenção do grau de Doutor em Engenharia Elétrica.

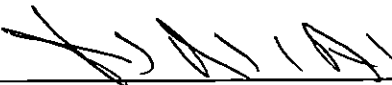
LUCIMAR MARIA FOSSATTI DE CARVALHO

Florianópolis, Abril 2007.

MODELOS DE APRENDIZAGEM PARA SISTEMAS NEURO-DIFUSOS

LUCIMAR MARIA FOSSATTI DE CARVALHO

Esta tese foi julgada adequada para a obtenção do Título de Doutor em Engenharia Elétrica, Área de Concentração em *Engenharia Biomédica*, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Santa Catarina.

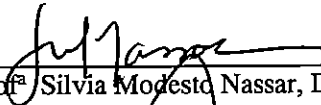


Prof^o Fernando Mendes de Azevedo, Dr.
Orientador

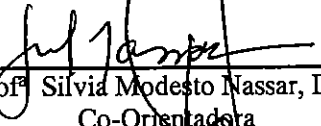
Prof^o Nelson Sadowski, Dr.

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica

Banca Examinadora:



Prof^ª Silvia Modesto Nassar, Dra.
Presidente



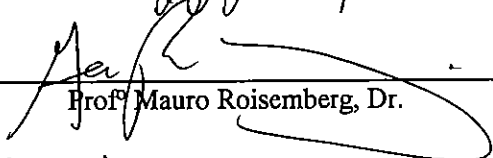
Prof^ª Silvia Modesto Nassar, Dra.
Co-Orientadora



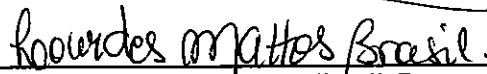
Prof^o Paulo Afonso Bracarense Costa, Dr.



Prof^ª Fernanda Isabel Marques Argoud, Dra.



Prof^o Mauro Roisemberg, Dr.



Prof^ª Lourdes Mattos Brasil, Dra.

Dedico

esta tese
à DEUS, pela força e
coragem com que me agraciou, dia após dia.

Anjo mensageiro de minha vida.

O SENHOR é meu pastor, nada me faltará.

Deitar-me faz em verdes pastos, guia-me mansamente
a águas tranquilas.

Refrigera a minha alma; guia-me pelas veredas da justiça,
por amor do seu nome.

Ainda que eu andasse pelo vale da sombra da morte,
não temeria mal algum,
porque *tu estás comigo*;
a tua vara e o teu cajado me consolam.

Preparas uma mesa perante mim na presença dos meus inimigos,
unges a minha cabeça com óleo,
o meu cálice transborda.

Certamente que a bondade e a misericórdia
me seguirão todos os dias da minha vida;
e, habitarei na casa do **SENHOR**
por longos dias.

SALMO 23

Reverencio

A memória dos meus avós

Luis Carlos e Miquelina Fossatti;
Eustáquio e Eudócia Rodrigues,
raízes de minha existência.

A minha querida mãe **Almeri**,
pelo incentivo para eu persistir SEMPRE.

Ao meu pai **Querino**, *in memóriam*,
que, mesmo nem sabendo o que eu tanto estudava,
estava sempre do meu lado, me apoiando em TUDO.

Enalteceram o meu caminho

Sérgio,

pela presença marcante em todos os momentos,
além da compreensão e paciência.

Priscilla e Patrícia,

que estiveram juntinho de mim, sempre !
Ora rindo, ora chorando, brincando, dizendo
a mãe só estuda...

Agradecimento com Louvor

Foi através da participação de vocês que esta tese concretizou-se.

A Universidade Federal de Santa Catarina, Ao Programa de Pós-Graduação em Engenharia Elétrica e, em especial, ao Instituto de Engenharia Biomédica pela acolhida recebida oportunizando crescimento profissional.

Aos professores membros da banca pelas preciosas sugestões.

À todos os professores do doutorado da UFSC, pela qualidade das disciplinas ministradas.

À Universidade de Passo Fundo (UPF), pelo apoio financeiro.

Ao Instituto de Ciências Exatas e Geociências (ICEG), pela oportunidade e incentivo.

Ao Curso de Ciência da Computação (CCC/UPF), e aos colegas do curso, pelas palavras de incentivo.

Dignifico

Meus ilustres orientadores,
que, através da sua sabedoria diária, nos preciosos encontros
me transmitiram a riqueza dos conhecimentos
que engrandeceram a minha alma
e, principalmente, me ensinaram que
quanto maior forem os obstáculos, maior será a glória de tê-los superado.

Meus agradecimentos e minha
profunda admiração pelo incomensurável apoio e dedicação:

Professor Dr. Fernando Mendes de Azevedo

Pela confiança em minha capacidade de desenvolver esta tese
me auxiliando nesta longa caminhada.

Sou e serei eternamente grata.

Professora Dra. Sílvia Modesto Nassar

Pela motivação que me passou na realização desta tese e
pelo exemplo de profissional,
será sempre uma inspiração na minha vida.

Sou e serei eternamente grata pela paciência, amizade e apoio.

Em especial, ao amigo Hugo José Teixeira de Carvalho,

médico neurologista,

pela amizade e colaboração.

Minha admiração pelo gesto impagável
na transmissão de conhecimentos da área neurológica.

Minha Eterna Gratidão

Professor Dr. Fernando Mendes de Azevedo

Professora Dra. Sílvia Modesto Nassar

Professor Dr. Paulo Afonso Bracarense Costa

Professora Dra. Fernanda Izabel Marques Argoud

Professor Dr. Mauro Roisemberg

Professora Dra. Lourdes Mattos Brasil

Banca que compôs esta tese,
tornando concreta mais uma etapa da minha formação acadêmica.

Presenças Preciosas na Minha Jornada

Agradeço de coração a todos que contribuíram para que o meu sonho se concretizasse.

Tudo deu certo, com o apoio, carinho, amizade e dedicação de vocês !

Cristiane e Nederson Koehler,
pelos primeiros incentivos e a primeira acolhida em vosso lar.

Magda e Neuza, pela paciência em me ensinar questões práticas da matemática.

Angelita de Ré, pela ajuda inestimável na revisão do sistema.

Orly Spanemberg, que me ensinou os primeiros passos no MATLAB.

Às amigas Tânia, Ionara, Mozen, Vânia, Dileta, Mariane, Bela e Rose, pelo carinho.

Aos meus preciosos bolsistas, Cândice e Lucas, pelo auxílio incansável no treinamento dos modelos.

Aos pacientes entrevistados, que dedicaram um tempo precioso de suas vidas, sou grata.

Ao amigo Paulo Rigon, pelas palavras de incentivo e acreditar no meu trabalho.

Aos meus colegas do doutorado no IEB pelos momentos de estudo:
Rosivete, Álvaro, Nazaré, Charles, Gonçalo, Patrícia, Rudimar, Andréa e ao Nelso. Em especial, agradeço pela amizade à todos os colegas do LEA1 e LEA2.

Ao amigo Masanao Ohira, pela disponibilidade no empréstimo de *softwares*.

Aos colegas do Hospital da Cidade de Passo Fundo, pelo companheirismo e amizade.

Michéli e Elisandra, estagiárias do CIPE, agradeço a ajuda na elaboração final dos grafos do conhecimento.

Mari, Angelita, Elaine e Emanuele, pela ajuda indispensável com os pacientes.

À Profª Maria Emilce, obrigada pela dedicação que dispensou na revisão ortográfica.

Ao colega Leandro Donin, agradeço pelas aulas de Delphi, imprescindível nos meus experimentos.

À minha Masoterapeuta Alexandra e a Neurolingüística Madalena,
que ouviam em silêncio as minhas lamentações.

Obrigada !

Agradecimento Especial

Carlos Javel do Vale, *in memóriam,*

**Meu eficiente Professor de Delphi,
juntos construímos os primeiros passos para a avaliação dos modelos,
porém o destino o enviou para outra dimensão,
não permitindo a ele
continuar até o fim.**

*“ ... os limites mais avançados do nosso ser mergulham em uma outra dimensão
de existência completamente diferente do mundo onde estamos, nós pertencemos a esta
outra dimensão...”*

Finalmente, agradeço com Deferência

Deus, meu Pai. **D**eus, minha mãe. **D**eus tudo em todos.

VOCÊ está presente onde quer que eu esteja.

Deus, eu me entrego a Você como uma oferenda de amor.

Que eu sempre seja uma expressão do Seu amor,
uma expressão da Sua verdade.

Que eu sempre pense de acordo com a Sua sabedoria.

Que eu sempre traga a energia do Seu Espírito
para cada situação,

de modo que Você possa tocar os corações
e as mentes de todas as pessoas.

Que eu me encha com a luz da Sua presença,
com o poder da Sua verdade.

Que eu possa viver de acordo com a fé
que Você deposita em mim.

Que eu satisfaça o objetivo que você traçou para mim.

Que eu aprenda a confiar sempre em Você, e
que a minha vida seja um exemplo luminoso da Sua sabedoria
e da Sua glória.

Que eu dignifique a Sua presença dentro de mim e
em minha vida vivendo
de acordo com o princípio do Seu amor.

Eu agradeço,
pois sei que Você irá me atender.

Por tudo aquilo que recebi
e que ainda está por vir,

Eu agradeço.

IYANLA VANZANT

PUBLICAÇÕES

Congressos e Simpósios

1. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese, SCHWALBERT, Michéli Adriana. Aplicabilidade das Operações Aritméticas Difusas *t-normas e t-conormas* no Diagnóstico Diferencial de Eventos Epilépticos e Não Epilépticos. In: *VII Workshop de Informática Médica*, Porto de Galinhas. Penambuco, Brasil, 2007.
2. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese, RECH, Ciciliana M. Z. Reconhecimento de Padrões aplicados aos Eventos Epilépticos. In: *13º Congresso de Pesquisa e Desenvolvimento em Saúde do Mercosul*. Porto Alegre: HCPA, Brasil, 2006.
3. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese. Neuro-Fuzzy Systems: Learning Models. In: *Proceedings of the IEEE International Conference On Computational Intelligence For Modelling Control And Automation*, (CIMCA-2005), Áustria, Viena, 2005.
4. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella; MONTEIRO, Lucas Lese. Uma Abordagem Neuro-Difusa de Classificação de Dados Utilizando a Operação Aritmética Min/Max. In: *V Workshop De Informática Médica*, Porto Alegre, Brasil, 2005.
5. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella. Modelo Híbrido de Aprendizagem utilizando Expressões Lingüísticas e

Numéricas no Auxílio ao Diagnóstico. *In: III Congresso Latino Americano de Engenharia Biomédica e IXI Congresso Brasileiro de Engenharia Biomédica*, 22 a 25 de Setembro, João Pessoa, Brasil, 2004.

6. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella. Proposta de um Modelo Híbrido de Aprendizagem utilizando diferentes T-normas e T-conormas. *In: IV Congresso Brasileiro de Computação e IV Workshop de Informática Aplicada à Saúde*, 08 a 12 de Outubro, Itajaí, Brasil, 2004.
7. CARVALHO, Lucimar Maria Fossatti; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella; MONTEIRO, Lucas Lese. Projeto de Redes Neurais Artificiais e Lógica Difusa no Auxílio ao Diagnóstico diferencial de Eventos Epilépticos e Não Epilépticos. *In: Congresso de Desenvolvimento em Saúde do Mercosul*, Porto Alegre, Brasil, 2004.
8. CARVALHO, Lucimar Maria Fossatti; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Projeto e Implementação de Arquiteturas SOMs's: Uma Aplicação na Área Neurológica. *In: Fourth Congress Of Logic Applied To Technologies*, 10 a 12 de Novembro, São Paulo, Brasil, 2003.
9. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Exploração de Bases de Dados Médicas utilizando Redes Neurais Artificiais auto-organizáveis. *In: Congresso Brasileiro de Engenharia Biomédica - CBEB*, 9 a 12 Setembro, São Paulo, Brasil, 2002.
10. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Implementação de técnicas utilizando Redes Neurais Artificiais auto-organizáveis no Diagnóstico

Médico. In: *VIII Jornadas Internacionales Ingeniería Clínica y Tecnología Médica*, 21 a 23 de Agosto, Buenos Aires, Argentina, 2002.

11. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Medical Data Base Exploration through Artificial Neural Networks. In: *4th International Conference On Enterprise Information - ICEIS*, 3 a 6 de Abril, p.455-459, Spain, 2002.
12. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Modelling Medical Diagnosis through Kononen Self-organizable Map. In: *4th Argentine Symposium On Artificial Intelligence*, 9 a 13 Setembro, p. 255-261, Buenos Aires, Argentina, 2002.
13. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. O processo de aprendizagem de Crises Convulsivas através de Redes Neurais Artificiais. *Anais do II Workshop de Informática Médica*, Gramado, Brasil, 2002.
14. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de. Redes Neurais Artificiais Auto-organizáveis utilizadas na Classificação de Crises Convulsivas. In: *Anais do Congresso Brasileiro de Informática na Saúde*, Natal, Brasil, 2002.
15. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; BRASIL, Ana Luisa Casado; DOZZA, Diego. Medical Data Base Exploration Through *Artificial Neural Networks*. In: *International Conference on Information and Knowledge Engineering*, 24 a 27 June, v. I, p. 209-215, Las Vegas, 2002.

Revistas

1. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese, RECH, Ciciliana M. Z. Reconhecimento de Padrões aplicados aos Eventos Epilépticos. Revista do HCPA, Hospital de Clínicas de POA, Indexada no LILACS, ISSN 0101-5575, v. 26, Supl. 1, p. 261-262, 2006.
2. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; DANI, Cândice Abella; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese; NASSAR, Sílvia Modesto, RECH, Ciciliana M. Z. Metodologia de Aprendizagem para Sistemas Neuro-Difusos. Revista Ciências Exatas e Naturais, Universidade Estadual Centro Oeste, ISSN 1518-0352, Ago-Dez, 2006.
3. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese. Avaliação de Modelos de Aprendizagem Aplicados em um Sistema de Diagnóstico Médico. Revista HCPA, Hospital de Clínicas de POA, Indexada no LILACS, ISSN 0101-5575, v. 25, n. 1, p. 245, 2005.
4. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; DANI, Cândice Abella; CARVALHO, Hugo José Teixeira de; MONTEIRO, Lucas Lese; NASSAR, Sílvia Modesto. Análise do Impacto de Diferentes Operações Aritméticas Aplicadas em um Sistema Neuro-Difuso. Revista Ciências Exatas e Naturais, Universidade Estadual Centro Oeste, ISSN 1518-0352, v. 6, n. 2, 2004.
5. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; NASSAR, Sílvia Modesto; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella. Análise de Modelos de Aprendizagem para Diagnosticar Eventos Epilépticos. Revista Brasileira de Neurologia, ISSN 0101-8469, UFRJ, v. 40, n. 2, p. 39-44, 2004.
6. CARVALHO, Lucimar Maria Fossatti de; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; NASSAR, Sílvia Modesto; DANI, Cândice

Abella; MONTEIRO, Lucas Lese. Projeto de Redes Neurais Artificiais e Lógica Difusa no Auxílio ao Diagnóstico Diferencial de Eventos Epilépticos e não Epilépticos. Revista do Hospital das Clínicas de Porto Alegre, Indexada no LILACS, ISSN 0101-5575, Porto Alegre, v. 1, 2004.

7. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella; DOZZA, Diego; BRASIL, Ana Luisa Casado. Uma Análise nas Técnicas de Aprendizagem utilizadas em Redes Neurais Artificiais. Revista Brasileira de Neurologia, ISSN 0101-8469, UFRJ, v. 39, Jan-Fev, p. 17-22, 2003.
8. CARVALHO, Lucimar Maria Fossatti de; DANI, Cândice Abella; BRASIL, Ana Luisa Casado; DOZZA, Diego; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; NASSAR, Sílvia Modesto. Análise de Técnicas de Reconhecimento de Padrões Aplicadas na Classificação de Crises Convulsivas. Revista do Hospital das Clínicas de Porto Alegre, Indexada no LILACS, ISSN 0101-5575, Porto Alegre, Brasil, v. 22, Novembro, p. 37-39, 2002.
9. CARVALHO, Lucimar Maria Fossatti de; NASSAR, Sílvia Modesto; AZEVEDO, Fernando Mendes de; CARVALHO, Hugo José Teixeira de; DANI, Cândice Abella; DOZZA, Diego; BRASIL, Ana Luisa Casado. Exploração da Bases de Dados Médicas através de Redes Neurais Artificiais. Revista Brasileira de Neurologia, ISSN 0101-8469, UFRJ, v. 38, Outubro, p. 32-37, 2002.

Resumo da Tese apresentada à UFSC como parte dos requisitos necessários para a obtenção do grau de Doutor em Engenharia Elétrica.

MODELOS DE APRENDIZAGEM PARA SISTEMAS NEURO-DIFUSOS

LUCIMAR MARIA FOSSATTI DE CARVALHO

Abril/2007

Orientador: Prof. Fernando Mendes de Azevedo, D. Sc.

Área de Concentração: Engenharia Biomédica

Palavras-chave: Redes Neurais Artificiais, Lógica Difusa, Sistemas Neuro-Difusos

Número de Páginas: 124

Esta tese tem por objetivo avaliar modelos de aprendizagem utilizando os conceitos Redes Neurais Artificiais (RNAs), Lógica Difusa (LD) e Sistemas Neuro-Difusos (SNDs). O maior interesse da pesquisa foi examinar a aplicabilidade das diferentes operações aritméticas difusas, implementadas através de dois tipos de neurônios difusos com o objetivo de obter taxas mais altas de acertos do SND. Sistemas híbridos combinam diferentes abordagens com a finalidade de integrar ambas as potencialidades.

A pesquisa fornece uma breve introdução às RNAs, seus princípios básicos de aprendizagem e os fundamentos que envolvem a representação do conhecimento difuso, descrevendo as arquiteturas e técnicas de inferência utilizadas nos sistemas híbridos neuro-difusos. A pesquisa apresenta uma revisão de importantes SNDs encontrados na literatura, como ANFIS e o NEFCLASS, procurando descrever a estrutura básica referente ao sistema de inferência difuso Mamdani, o qual foi utilizado na tese.

Os resultados mostram que as operações aritméticas difusas Soma/Produto de Einstein OU(E), implementadas com o neurônio difuso proposto por Kwan-Cai, obtiveram os melhores índices de acertos do SND. Os resultados foram analisados considerando os indicadores quantitativos de modelagem. A área de domínio utilizada para avaliar os modelos de aprendizagem foi no diagnóstico de eventos epiléticos (EEs) e eventos não epiléticos (ENEs).

Thesis abstract presented to UFSC as a partial fulfillment of the requirements for the degree of Doctor in Electrical Engineering.

LEARNING MODELS FOR NEURO-FUZZY SYSTEMS

LUCIMAR MARIA FOSSATTI DE CARVALHO

April/2007

Advisor: Prof. Fernando Mendes de Azevedo, D. Sc.

Area of Concentration: Biomedical Engineering

Keywords: Artificial Neural Networks, Fuzzy Logic and Fuzzy Neural Networks

Number of Pages: 124

The goal of this thesis is to appraise learning models by integrating the concepts among Artificial Neural Networks (ANNs), Fuzzy Logic (FL) and Neuro-Fuzzy Systems (NFSs). The main interest of the research was to examine the applicability of the different fuzzy arithmetical operations, implemented by means of two types of fuzzy neurons to obtain the highest rates of NFS performance. Hybrid systems combine different approaches in order to take advantage of their respective strengths.

The research supplies a brief introduction to ANNs, their basic learning principles and the fundamentals which involve the representation of fuzzy knowledge by describing the architectures and inference techniques utilized in the hybrid systems neuro-fuzzy. The research presents a review of important NFSs which are found in literature, such as ANFIS and NEFCLASS, seeking to describe the basic structure referring to the Mandani fuzzy inference system, which was utilized in the thesis.

The results show that the fuzzy arithmetical operations of Einstein's Sum/Product OR(AND) implemented with the fuzzy neuron proposed by Kwan-Cai has obtained the highest rates of NFS performance. The results have been analyzed by taking into account the modeling quantitative indicators. The domain area, which is used to appraise learning models, was in the diagnosis of Paroxysmal Events involving epileptic events (EEs) and non epileptic events (NEEs).

Sumário

SUMÁRIO	i
LISTA DE FIGURAS	iv
LISTA DE QUADROS	v
LISTA DE SIGLAS	vi
1- INTRODUÇÃO.....	1
1.1- Objetivos da Pesquisa.....	4
1.1.1-Objetivo Geral.....	4
1.1.2-Objetivos Específicos.....	5
1.2- Justificativa e Importância da Pesquisa	5
1.3- Organização da Tese.....	5
2- FUNDAMENTAÇÃO TEÓRICA.....	7
2.1- Redes Neurais Artificiais.....	7
2.1.1-Considerações Iniciais	7
2.1.2-Modelo de Neurônios.....	7
2.1.3-Modelos de Redes Neurais Artificiais.....	10
2.1.4-Formas de Aprendizagem nas RNAs	12
2.1.5-Algoritmo de Retropropagação.....	17
2.1.6-Considerações Finais	19
2.2- Lógica Difusa	20
2.2.1-Considerações Iniciais	20
2.2.2-Conjuntos Difusos.....	21
2.2.3-Operações com Conjuntos Difusos.....	23
2.2.4-Base de Regras.....	24
2.2.5-Sistemas de Inferência Difusos.....	25
2.2.6-Considerações Finais	27
2.3- Sistemas Neuro-Difusos	28
2.3.1-Considerações Iniciais	28
2.3.2-Tipos de SNDs	29
2.3.3-Projeto de Sistemas Neuro-Difusos	31
2.3.4-Neurônios Difusos	32
2.3.5-Arquiteturas de Sistemas Neuro-Difusos.....	36
2.3.6-Considerações Finais	48
3- MODELOS DE APRENDIZAGEM.....	49
3.1- Considerações Iniciais	49
3.2- Metodologias de Aprendizagem.....	50
3.2.1-Aprendizagem com o Neurônio Difuso Hirota-Pedrycz (1992, 1994).....	51
3.2.2-Aprendizagem com o Neurônio Difuso Kwan-Cai (1994).....	52
3.2.3-Aprendizagem com as Operações Aritméticas <i>Max</i> e <i>Min</i> (Nauck, 1994).....	54
3.2.4-Restrições Referente às Operações Aritméticas <i>Max/Min</i>	56
3.3- Originalidade da Pesquisa	56
3.4- Considerações Finais.....	57

4- DOMÍNIO DE APLICAÇÃO	58
4.1- Considerações Iniciais	58
4.2- Classificação dos Eventos Epilépticos.....	59
4.2.1-International League Against Epilepsy (ILAE, 1981)	60
4.2.2-Características dos Eventos Epilépticos.....	60
4.3- Sistematização dos Eventos Não Epilépticos	63
4.3.1-Classificação dos Eventos Não Epilépticos	63
4.3.2-Características dos Eventos Não Epilépticos.....	65
4.4- Considerações Finais	66
5- IMPLEMENTAÇÃO DOS MODELOS DE APRENDIZAGEM.....	68
5.1- Considerações Iniciais	68
5.2- Implementação da Base de Conhecimento.....	68
5.2.1-Aquisição e Representação do Conhecimento.....	69
5.2.2-Base de Regras	70
5.2.3-Bases de Treinamento e de Teste	71
5.2.4-Variáveis de Entrada	71
5.2.5-Variável de Saída.....	74
5.3- Arquitetura do SisND/EPs.....	75
5.3.1-Modelagem do Ambiente.....	75
5.3.2-Interface	78
5.4- Avaliação dos Modelos de Aprendizagem	84
5.4.1-Avaliação do Desempenho dos Modelos (Nassar, 1998).....	84
5.4.2-Validação Interna da Base de Conhecimentos (Nassar, 1998).....	85
5.5- Considerações Finais	86
6- RESULTADOS E DISCUSSÃO	88
6.1- Considerações Iniciais	88
6.2- Aprendizagem dos Modelos	88
6.2.1-NEFCLASS.....	89
6.2.2-Hirota-Pedrycz	89
6.2.3-Kwan-Cai	93
6.2.4-Desempenho das Arquiteturas NEFCLASS e do SisND/EPs.....	98
6.3- Avaliação de Desempenho das Arquiteturas NEFCLASS e do SisND/EPs.....	99
6.3.1-Teste de Hipótese das Arquiteturas NEFCLASS e do SisND/EPs	99
6.4- Validação Interna	100
6.4.1-Neurônio Difuso Kwan-Cai	100
6.4.2-Neurônio Difuso Hirota-Pedrycz	101
6.4.3-Arquitetura NEFCLASS - Operações Aritméticas Difusas Padrões <i>Max/Min</i>	102
6.4.4-Sensibilidade e Valor Preditivo Positivo	103
6.4.5-Especificidade e Valor Preditivo Negativo	104
6.5- Comportamento da Aprendizagem.....	105
6.5.1-NEFCLASS e do SisND/EPs.....	105
6.5.2-Sensibilidade, Especificidade, Valor Preditivo Positivo e Negativo	107
6.6- Considerações Finais	108
7- CONCLUSÕES.....	110

REFERÊNCIAS	113
ANEXOS	125
APÊNDICES	130

Lista de Figuras

Figura 2.1.1 – Neurônio biológico (adaptado de Haykin (2001))	08
Figura 2.1.2 – Modelo de neurônio McCulloch e Pitts (adaptado de Freeman e Skapura (1991))	09
Figura 2.1.3 – Arquitetura de uma RNA <i>Perceptron</i> Multicamadas	11
Figura 2.2.1 – Representação da função de pertinência (adaptado de Rezende et al. (2003))	22
Figura 2.2.2 – Operações com conjuntos difusos	23
Figura 2.2.3 – SID <i>Mamdani</i> (adaptado de Gorzalczany (2002))	25
Figura 2.2.4 – SID <i>Mamdani</i> (adaptado de Melin e Castillo (2005))	27
Figura 2.3.1 – Primeiro SND (adaptado de Fullér (2000))	29
Figura 2.3.2 – Segundo SND (adaptado de Fullér (2000))	30
Figura 2.3.3 – Modelo de SND (adaptado de Kartalopoulos (1996))	30
Figura 2.3.4 – Neurônio difuso E (adaptado de Hirota e Pedrycz (1992, 1994))	32
Figura 2.3.5 – Neurônio difuso OU (adaptado de Hirota e Pedrycz (1992, 1994))	33
Figura 2.3.6 – Neurônio difuso E (adaptado de Kwan e Cai (1994))	34
Figura 2.3.7 – Neurônio difuso OU (adaptado de Kwan e Cai (1994))	35
Figura 2.3.8 – Arquitetura ANFIS (adaptada de Jang, Sun e Mizutani (1997))	36
Figura 2.3.9 – Operadores lógicos E/OU (adaptado de Mitra e Pal (1992))	39
Figura 2.3.10 – SND FuNe-I (adaptado de Halgamuge e Glesner (1994))	40
Figura 2.3.11 – SND com neurônios OU (adaptado de BRASIL (1999))	41
Figura 2.3.12 – Inferência difusa E/OU e OU (adaptado de Pedrycz (1993))	44
Figura 2.3.13 – Arquitetura NEFCLASS (adaptada de Nauck e Kruse (1995, 1996))	46
Figura 4.1.1 – Classificação dos eventos paroxísticos (adaptado de Costa et al. (1998))	63
Figura 5.1.1 – Conjunto difuso das variáveis lingüísticas	72
Figura 5.1.2 – Variável lingüística “tempo de duração”	73
Figura 5.1.3 – Arquitetura do SisND/EPs	76
Figura 5.1.4 – Parâmetros de configuração	79
Figura 5.1.5 – Neurônios difusos e as operações aritméticas difusas	80
Figura 5.1.6 – Inclusão e alteração de regras	80
Figura 5.1.7 – Fase de conferência do SisND/EPs	81
Figura 5.1.8 – Tela de ajuda ao usuário	82
Figura 5.1.9 – Fase de treinamento e de teste	82
Figura 5.1.10 – Gráfico em movimento	83
Figura 6.1.1 – EMQ – <i>Max/Min</i> para a arquitetura NEFCLASS	89
Figura 6.1.2 – EMQ – Soma/Produto Algébrico OU(E) para o neurônio difuso Hirota-Pedrycz	89
Figura 6.1.3 – EMQ – Soma/Produto de Hamacher OU(E) para o neurônio difuso Hirota-Pedrycz	90
Figura 6.1.4 – EMQ – Soma/Produto de Einstein OU(E) para o neurônio difuso Hirota-Pedrycz	90
Figura 6.1.5 – EMQ – <i>Max/Min</i> para o neurônio difuso Hirota-Pedrycz	91
Figura 6.1.6 – EMQ da arquitetura NEFCLASS e do neurônio difuso Hirota-Pedrycz	91
Figura 6.1.7 – EMQ – Soma/Produto Algébrico OU(E) para o neurônio difuso Kwan-Cai	94
Figura 6.1.8 – EMQ – Soma/Produto de Hamacher OU(E) para o neurônio difuso Kwan-Cai	94
Figura 6.1.9 – EMQ – Soma/Produto de Einstein OU(E) para o neurônio difuso Kwan-Cai	95
Figura 6.1.10 – EMQ – <i>Max/Min</i> para o neurônio difuso Kwan-Cai	95
Figura 6.1.11 – EMQ da arquitetura NEFCLASS e do neurônio difuso Kwan-Cai	96

Lista de Quadros

Quadro 3.1.1 – Arquiteturas de SND e as operações aritméticas difusas.....	49
Quadro 5.1.1 – Teste de hipóteses	84
Quadro 5.1.2 – Medidas de validade interna dos SNDs	85
Quadro 6.1.1 – Desempenho das arquiteturas NEFCLASS e Hirota-Pedrycz: fase de treinamento .	93
Quadro 6.1.2 – Desempenho das arquiteturas NEFCLASS e Kwan-Cai: fase de treinamento	97
Quadro 6.1.3 – Comparação de desempenho das arquiteturas NEFCLASS e do SisND/EPs	98
Quadro 6.1.4 – Teste de Hipótese das arquiteturas NEFCLASS e do SisND/EPs	99
Quadro 6.1.5 – Validação interna – neurônio difuso Kwan-Cai	100
Quadro 6.1.6 – Validação interna - neurônio difuso Hirota-Pedrycz	101
Quadro 6.1.7 – Validação interna - operações aritméticas difusas padrões <i>Max/Min</i>	102
Quadro 6.1.8 – Validação interna - sensibilidade e valor preditivo positivo	103
Quadro 6.1.9 – Validação interna - especificidade e valor preditivo negativo	104
Quadro 6.1.10 – Comportamento da aprendizagem – NEFCLASS e do SisND/EPs	106
Quadro 6.1.11 – Aprendizagem - sensibilidade, especificidade, valor preditivo positivo e negativo	107

Lista de Siglas

As siglas estão assim distribuídas:

AG - algoritmos Genéticos

ANFIS - *Adaptative Network based Fuzzy Inference Systems*

CDs - Conjuntos Difusos

CE - Computação Evolucionária

CGTC - Crise Generalizada Tônico-Clônica

CHs - Crises Histéricas

CID - Cadastro Internacional de Doenças

CPC - Crise Parcial Complexa

CPS - Crise Parcial Simples

EEs - Eventos Epilépticos

EEC - Eletrocardiograma

EEG - Eletroencefalograma

EMQ - Erro Médio Quadrático

ENEs - Eventos Não Epilépticos

EPs - Eventos Paroxísticos

IA - Inteligência Artificial

IC - Inteligência Computacional

ILAE - *International League Against Epilepsy*

LD - Lógica Difusa

ML - *Machine Learning*

MCP - McCulloch Pitts

MLPs - *MultiLayers Perceptrons*

NEFCLASS - *Neuro-Fuzzy Classification*

NEFCON - *Neuro-Fuzzy Controller*

RNAs - Redes Neurais Artificiais

RNBs - Redes Neurais Biológicas

SIDs - Sistemas de Inferência Difusos

SDNs - Sistemas Difuso-Neurais

SIHs - Sistemas Inteligentes Híbridos

SisND/EPs - Sistema Neuro-Difuso para Diagnóstico de Eventos Paroxísticos

SNC - Sistema Nervoso Central

SNDs - Sistemas Neuro-Difusos

SQL - *Structured Query Language*

TSK - *Takagi Sugeno Kang*

Capítulo 1

Introdução

Entre as diversas áreas de pesquisa envolvendo a Ciência da Computação destacam-se a investigação e a simulação de modelos de aprendizagem. Um dos grandes desafios encontrados na pesquisa computacional é delinear esses modelos de aprendizagem de forma a criar máquinas inteligentes.

Segundo Luger (2004), o principal objetivo da aprendizagem de máquina é o estudo e a compreensão de modelos de aprendizagem por meio de algoritmos aplicados em sistemas computacionais. Qualquer sistema computacional que deseje trabalhar a inteligência deve possuir a habilidade de aprender, isto é, a capacidade de adquirir novas formas de conhecimento. O autor enfatiza que a aprendizagem envolve a generalização a partir da experiência e que o desempenho do sistema computadorizado deve melhorar, não apenas na repetição da mesma tarefa, mas também em tarefas similares do domínio.

Essas características são complexas e apresentam dificuldades na automatização visto que envolvem diferentes habilidades pertinentes ao raciocínio humano, diferentes formas do raciocínio (lógica), intuição, raciocínio em situações de incerteza, pensamento e crenças as quais devem fazer parte do processo de aprendizagem de máquina.

A aprendizagem de máquina ocorre em dois grandes paradigmas: a aprendizagem com um professor ou supervisionada e a aprendizagem sem um professor ou não-supervisionada (Haykin, 2001).

Todos esses aspectos relacionados com a cognição humana, aplicados na resolução de problemas, conduzem ao projeto de programas computacionais envolvendo a área da Inteligência Artificial (IA).

Na literatura existem diferentes definições para IA, que, de uma forma ou outra, procuram deixar mais claros alguns conceitos. Entre eles, citam-se:

Russel e Norvig (2004) descrevem a IA utilizando duas formas complementares:

(1) relacionada com o raciocínio, envolvendo questões referentes ao senso empírico, formulação de hipóteses e experimentos;

(2) direcionada ao comportamento, utilizando fundamentos matemáticos.

Uma definição sugerida por Luger e Stubblefield (1993) é da IA como o ramo da Ciência da Computação que se ocupa da automatização do comportamento inteligente.

Feigenbaum e McCorduck (1983) assim se expressam: "se pudermos imaginar um dispositivo capaz de colecionar, de selecionar entre, de compreender, de distinguir e de saber, então temos IA".

Com base nessas definições, pode-se concluir que as pesquisas na área da IA integram tanto questões envolvendo o raciocínio como questões relacionadas com o comportamento humano¹.

O universo das pesquisas apresenta uma tendência a ser modificado freqüentemente, pois a realidade é dinâmica, os métodos e as técnicas evoluem rapidamente, contribuindo para o surgimento de novos paradigmas de pesquisa. Os sistemas envolvendo técnicas tradicionais de IA estão dando lugar aos sistemas que utilizam Inteligência Computacional (IC), referenciada na literatura como *soft computing* (Melin e Castillo, 2005), que consiste de vários paradigmas computacionais, incluindo Lógica Difusa (LD), Redes Neurais Artificiais (RNAs) e algoritmos Genéticos (AG).

No centro da metodologia da IC estão as técnicas de IA, tais como a LD, Neurocomputação, Computação Evolucionária, Computação Probabilística, Computação Caótica e a Aprendizagem de Máquina (Zadeh, 2006).

A IC busca, por meio de técnicas inspiradas na natureza, o desenvolvimento de sistemas inteligentes que possam simular aspectos do comportamento humano, tais como aprendizado, percepção, raciocínio, evolução e adaptação (Velasco e Pacheco, 2006).

Segundo Rutkowska (2002), a IC fornece conceitos e técnicas básicas para a tomada de decisão e automação do raciocínio em domínios com imprecisão e incerteza. Esses sistemas buscam a representação do conhecimento inexato por meio de modelos híbridos de aprendizagem.

Isso se justifica, uma vez que no mundo real observa-se a presença de diferentes formas de inexatidão, tais como informações incertas e imprecisas, simultaneamente. Em razão disso, ao se projetar um sistema envolvendo o trabalho conjunto destas técnicas de IA, os resultados obtidos com o sistema poderão simular a forma do raciocínio humano na resolução de problemas complexos.

Existem diferentes autores que modelam a incerteza segundo diferentes abordagens, como imprecisão, vagueza, ambigüidade, conflito, ignorância, etc. Para cada tipo de abordagem existe um modelo formal de representação, que varia segundo diferentes teorias. Uma teoria pode ser apropriada para representar uma determinada informação, porém poderá ser inapropriada para representar outro tipo de informação.

¹ Outras referências relacionadas com a área da IA são encontradas em Aaai (2007), Aisb (2007), Auai (2007), Sigarr (2007) e Mit (2007).

A informação imprecisa e/ou vaga pode ser modelada utilizando-se a teoria dos CDs, que é apropriada para medir o “significado” da informação. Dessa forma, um elemento pertence com maior ou menor intensidade a uma determinada classe, utilizando grau de verdade (Zadeh, 1965).

Zadeh também propõe a teoria de possibilidades para modelar as informações vagas (Zadeh, 1978). Segundo o autor, a imprecisão é um componente da vagueza, que está associada à dificuldade de fazer distinções precisas do mundo real, ou seja, dificuldade de estabelecer limites precisos bem definidos (Klir e Folger, 1988).

A informação probabilista que mede a aleatoriedade pode ser modelada pela teoria de probabilidades ou através da teoria de evidências, que atribui a cada elemento um grau numérico de crença².

Costa (1999) descreve de forma clara as idéias de Smithson (1989), referenciando a "incerteza" no contexto da "ignorância". O autor descreve a "ignorância" como uma medida que envolve o "erro" e a "irrelevância". A "irrelevância" (ato de ignorar) procura identificar a não-topicalidade, o tabu e a não-decidibilidade das informações. O "erro" (estado de ignorância) está dividido em distorção e incompletude das informações. A distorção envolve a confusão e a inacurácia das informações. Na incompletude das informações, encontram-se a incerteza e a ausência de informações. Relacionados com a incerteza, encontram-se os termos probabilísticos e ambíguos. E, finalmente, os termos vagos se dividem em difusividade *fuzzyness* e a não-especificidade³.

Uma das vantagens da neurocomputação que trabalha com as RNAs encontra-se na sua capacidade de aprendizagem em domínios com conhecimento incerto. Por outro lado, a vantagem da LD encontra-se na capacidade de modelagem das informações imprecisas.

As pesquisas mais recentes em SNDs apontam para a determinação automática do número de termos lingüísticos para a *fuzzi ficação* de cada variável (Rezende et al., 2003).

Os sistemas híbridos envolvendo as RNAs e a LD possuem um grande potencial para modelar domínios com conhecimento impreciso e incerto. Assim, investigou-se modelos de aprendizagem aplicados aos Sistemas Neuro-Difusos (SNDs) por meio de diferentes operações aritméticas difusas e diferentes tipos de neurônios difusos modelando a "imprecisão" presente nas entradas da RNA. Os resultados da pesquisa, após o processo de avaliação e validação, foram aplicados aos SNDs propostos com o objetivo de desenvolver sistemas com melhores percentuais de acertos. Estes modelos de

² A teoria de evidências é conhecida como teoria das evidências de Dempster-Shafer (Shafer, 1976).

³ Para uma compreensão mais detalhada das diferentes naturezas relacionadas com as medidas de incerteza consultar a obra original de Smithson (1989).

aprendizagem foram avaliados no diagnóstico dos *Eventos Epilépticos (EEs)* e dos *Eventos não Epilépticos (ENEs)*.

Na literatura são encontradas diferentes arquiteturas híbridas aplicadas no desenvolvimento de SNDs. As arquiteturas mais encontradas concentram-se na utilização das operações aritméticas difusas padrões *max/min* e da operação aritmética difusa produto algébrico⁴ no processo de inferência. Isso leva à investigação de modelos de aprendizagem em SNDs explorando diferentes operações aritméticas difusas, entre as quais: Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E) e Soma/Produto de Einstein OU(E), além das operações aritméticas difusas padrões *max/min*.

Pedrycz e Rocha (1993) justificam que a seleção adequada das operações aritméticas difusas pode refletir na redução da complexidade computacional e na otimização da representação do conhecimento pelos neurônios difusos.

Na área da Engenharia Biomédica, existem várias pesquisas que utilizam RNAs e LD como ferramentas. Entre elas, pode-se citar:

- Pesquisas na área de Sistemas Especialistas *Fuzzy* (De Azevedo et al., 2004; Polat et al., 2007);
- Pesquisas na área de RNAs (Argoud, 2001; Dua et al., 2004; Hahne, 2005; Iaione, 2003);
- Pesquisas na área de Sistemas Especialistas Híbridos (Appel, 2004; Brasil, De Azevedo e Barreto, 2001; Javad, Karim e Majid, 2003; Lim et al.; 2005; Vianna, 2000).

Desta forma, faz-se o seguinte questionamento: como a modelagem por meio da LD pode contribuir para que a aprendizagem das RNAs torne o SND mais eficiente?

Nesta tese os modelos de aprendizagem para SNDs propostos, são avaliados na área da neurologia, e, dentro desta área, no diagnóstico diferencial de *EEs* e os *ENEs*.

1.1 Objetivos da Pesquisa

1.1.1 Objetivo Geral

Avaliar modelos de aprendizagem para Sistemas Neuro-Difusos explorando diferentes operações aritméticas difusas na área médica.

⁴ Algumas referências bibliográficas, que utilizam as operações aritméticas difusas padrões, encontram-se no Apêndice 1.

1.1.2 Objetivos Específicos

Os objetivos específicos da tese são:

- explorar modelos de neurônios difusos;
- investigar diferentes operações aritméticas difusas;
- avaliar o desempenho dos modelos de aprendizagem para SNDs no diagnóstico dos *EEs* e *ENEs*;
- validar os modelos de aprendizagem para SNDs no diagnóstico dos *EEs* e *ENEs*;
- identificar modelos de aprendizagem com maior desempenho.

1.2 Justificativa e Importância da Pesquisa

A maioria dos algoritmos de aprendizagem de máquina utilizados computacionalmente para classificação de padrões, os quais utilizam RNAs e LD, fazem uso das operações aritméticas difusas padrões *min/max*. Na tentativa de obter melhores resultados de acertos nos modelos de aprendizagem e, mesmo, para a utilização destas informações conjuntamente com outras técnicas, faz-se necessária a investigação de novas estruturas e novas técnicas de modelos de aprendizagem.

Portanto, a investigação de diferentes operações aritméticas difusas, juntamente com a exploração de diferentes tipos de neurônios difusos, indica uma possibilidade de se obterem melhores resultados.

1.3 Organização da Tese

Esta tese está organizada em sete capítulos, acrescidos de quatro anexos e seis apêndices. O segundo capítulo fornece a fundamentação teórica dos tópicos relacionados com as RNAs, LD e SNDs. São feitas uma introdução às RNAs, a descrição dos princípios básicos relacionados com a LD e uma revisão das principais etapas utilizadas no projeto de SNDs.

No terceiro capítulo são apresentados os modelos de aprendizagem, uma descrição das arquiteturas de SNDs com suas respectivas operações aritméticas difusas, incluindo os diferentes tipos de neurônios difusos. Os tópicos do quarto capítulo estão relacionados com o domínio de aplicação utilizado para avaliar a metodologia e descrevem o diagnóstico médico em epilepsia.

O quinto capítulo mostra a implementação dos modelos de aprendizagem, definindo o processo de aquisição do conhecimento, base de conhecimento, juntamente com as variáveis de entrada e saída do SND descrevendo a arquitetura do *Sistema Neuro-Difuso para Diagnóstico de Eventos Paroxísticos* (SisND/EPs). Faz-se uma descrição da metodologia de avaliação dos modelos de aprendizagem explorados.

O sexto capítulo contém os resultados da simulação, incluindo uma comparação com as diferentes operações aritméticas difusas. Finalmente, o capítulo 7 apresenta as conclusões e sugestões para pesquisas futuras.

Na seqüência são descritas as referências consultadas para a realização da tese. Nos anexos encontram-se o Formulário de Coleta de Dados do Diagnóstico em Eventos não Epilépticos (Anexo 1); Coleta de Dados do Diagnóstico em Eventos Epilépticos (Anexo 2)⁵; Termo de Consentimento Livre e Esclarecido (Anexo 3) e Resolução do Comitê de Ética em Pesquisa da Universidade de Passo Fundo, com parecer favorável à pesquisa (Anexo 4). Nos Apêndices encontram-se a base de treinamento (Apêndice 1); base de teste (Apêndice 2); lista de variáveis de entrada (Apêndice 3); algoritmos de aprendizagem (Apêndice 4); grafos do conhecimento (Apêndice 5) e glossário (Apêndice 6).

⁵ Ambos os formulários de coleta de dados (Anexo 1 e 2) recebem valores contínuos baseados na escala de Likert (1952).

Capítulo 2

Fundamentação Teórica

Neste capítulo são descritos resumidamente os principais tópicos da teoria envolvendo RNAs, LD e SNDs. Esses são os conceitos necessários para a realização desta tese.

2.1 Redes Neurais Artificiais

Esta seção fornece uma breve introdução às RNAs. Inicialmente, é apresentado o modelo de neurônios, incluindo o neurônio biológico e o artificial. Faz-se uma descrição das formas de aprendizagem: hebbiana, regra delta e regra delta generalizada. Para finalizar, são apresentados dois diferentes modelos de RNAs, *Perceptrons* e o *MultiLayer Perceptrons* (MLPs).

2.1.1 Considerações Iniciais

RNAs são sistemas paralelos e distribuídos compostos por unidades de processamento (nós) que calculam determinadas funções matemáticas. Estas unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões. Na maioria dos modelos estas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da RNA (Braga, Carvalho e Ludermir, 2000).

As pesquisas com as RNAs iniciaram no campo da neurobiologia matemática. Os modelos recentes estão caminhando na direção da área estatística, principalmente em aplicações nas áreas do comércio, gerenciamento de informações, *marketing*, sistemas de apoio à decisão, sistemas de informação na produção industrial e engenharia (Vellido, 2002).

2.1.2 Modelo de Neurônios

O neurônio artificial, criado com base no neurônio biológico, representa a unidade principal no processamento das RNAs. Percebe-se que ambos os neurônios compõem uma grande RNA

organizada de diversas formas, apresentando como característica principal o processamento paralelo e distribuído.

· *O Neurônio Biológico*

O Sistema Nervoso Central (SNC) contém mais de cem bilhões de neurônios, compostos por três partes principais: os dendritos, o soma ou corpo celular, descrito como o corpo principal do neurônio, e um axônio único, que se estende do soma para o nervo periférico percorrendo os dendritos (Guyton, 2006).

A Figura 2.1.1 mostra um neurônio típico encontrado no córtex cerebral motor.

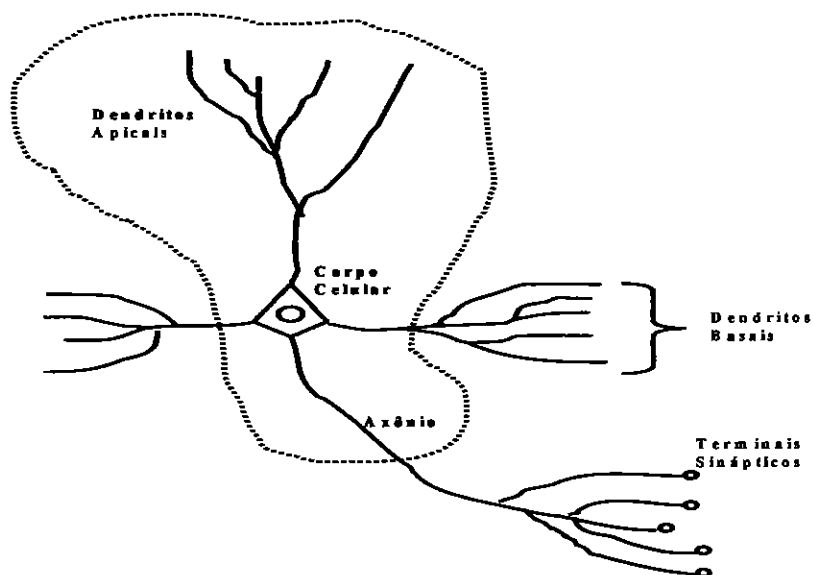


Figura 2.1.1 - Neurônio biológico (adaptado de Haykin (2001))

Os sinais chegam ao neurônio por meio das sinapses, passando pelos dendritos. As junções entre o axônio de uma célula e o dendrito de outra célula constituem as sinapses. O sinal de saída trafega por meio do axônio, o qual possui muitos ramos distintos que atingem outras partes do SNC. Este sinal normalmente passa do axônio do neurônio para os dendritos de outro neurônio, contribuindo, dessa forma, para o desempenho das diferentes funções nervosas. As sinapses constituem estímulos entre os neurônios e determinam a direção em que os sinais nervosos se distribuem pelo SNC constituindo uma ação seletiva. É através das sinapses que os neurônios se unem para formar as RNAs neurais.

Uma das funções mais importantes do SNC é o processamento das informações. O SNC recebe milhares de informações dos diferentes receptores sensoriais; combinando-as com as informações

armazenadas, obtidas pelos mecanismos de aprendizagem, desta forma o SNC produzirá uma resposta a ser dada pelo corpo.

Uma das funções mais importantes do SND é o processamento das informações. O SNC recebe milhares de informações dos diferentes receptores sensoriais; combinando-os com as informações armazenadas, obtidas pelos mecanismos de aprendizagem, desta forma o SNC produzirá uma resposta a ser dada pelo corpo.

· *O Neurônio Artificial*

Um neurônio artificial modela os axônios e os dendritos por meio de conexões sinápticas utilizando pesos de ajuste, de uma forma adequada, e então conecta esses modelos de uma maneira altamente paralela, de forma similar ao que ocorre realmente nas redes nervosas biológicas. Este trabalho possibilita um mecanismo de processamento complexo, com capacidade de aprendizagem (Passold, 2003).

O primeiro modelo artificial de um neurônio biológico foi idealizado pelo fisiologista e psiquiatra Warren McCulloch, juntamente com seu aluno, o matemático Walter Pitts (McCulloch e Pitts, 1943), conhecido como modelo McCulloch Pitts (MCP). No modelo MCP, o funcionamento do neurônio possui características similares às de um circuito binário (Figura 2.1.2).

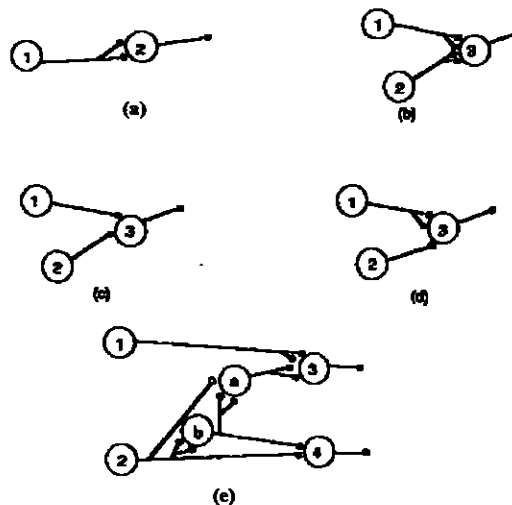


Figura 2.1.2- Modelo de neurônio McCulloch e Pitts (adaptado de Freeman e Skapura (1991))

Na Figura 2.1.2 o modelo MCP, baseado no circuito neural do sistema nervoso central, é definido utilizando a lógica proposicional. Os círculos grandes representam o corpo celular seguido dos dendritos; os círculos menores representam as conexões sinápticas: excitatórias (círculo fechado)

e inibitórias (círculo aberto). A Figura 2.1.2 (a) ilustra a precessão; (b) intersecção; (c) união ; (d) negação e (e) uma combinação geral do modelo.

2.1.3 Modelos de Redes Neurais Artificiais

São vários os modelos de RNAs encontrados na literatura, cada um com sua própria arquitetura, diferentes processos de aprendizagem e capacidade de desempenho distinta dos demais.

A proposta metodológica desta tese faz uso de RNAs tipo *Multilayer Perceptrons* (MLPs), o que implica que os modelos *Perceptrons* e o MLPs são os únicos modelos de RNAs aqui descritos.

· *Perceptrons*

O modelo *Perceptron* foi proposto por Frank Rosenblatt, em 1958. Sua topologia de RNA pode apresentar uma única camada de neurônios com um ou mais neurônios de saída (Rosenblatt, 1958).

Dessa forma, um neurônio i com p entradas do tipo $w_{ij}u_j$ apresenta a função net_i^p definida por:

$$net_i^p = \sum_j w_{ij}u_j + \theta_i \quad (2.1.1)$$

onde:

net_i^p = entrada da RNA do neurônio i para o padrão p ;

w_{ij} = intensidade da conexão do neurônio j da camada de entrada para o neurônio i da camada de saída;

u_j = valor de entrada do neurônio j ;

θ_i = representa um valor de polarização (*bias*) que pode assumir valores entre -1 até +1.

A saída da RNA é definida por:

$$y_i^p = \varphi(net_i^p) \quad (2.1.2)$$

onde:

y_i^p = saída do neurônio i para o padrão p ;

$\varphi(net_i^p)$ = função de ativação avaliada na entrada da RNA net_i^p atingido pelo neurônio i da camada de saída da RNA, no caso do *Perceptron* é uma saída não-linear.

A proposta de Rosenblatt em relação ao modelo de McCulloch e Pitts apresenta as seguintes características (Kovacs, 1996): (1) pode ser estendido para assumir valores contínuos, em oposição aos valores binários; (2) introduz uma lei de treinamento para os algoritmos *Perceptrons*.

Em 1969, Minsky e Pappert escreveram um livro onde foram descritas algumas limitações básicas no *Perceptron* de duas camadas, principalmente pela impossibilidade de implementar o *ou exclusivo*, sem a utilização da camada intermediária (Minsky e Pappert, 1969). Em sua obra, Minsky

e Pappert afirmaram que não acreditavam na possibilidade de desenvolver um algoritmo para resolver tal problema. Mais tarde, no entanto, foi provado o contrário, o que deu origem às RNAs MLPs, treinadas pelo algoritmo *backpropagation*⁶, descritas a seguir.

· *Multilayer Perceptrons (MLPs)*

As aplicações envolvendo RNAs MLPs encontram-se em diferentes áreas, desempenhando tarefas tais como processamento de sinais e imagens, classificação de padrões, controle, otimização, etc.

As RNAs MLPs apresentam pelo menos uma camada intermediária e podem trabalhar com dados que não são linearmente separáveis (Rumelhart, Hinton e Williams, 1986). As camadas de neurônios se intercomunicam, isto é, os dados fluem da camada de entrada para a camada intermediária e, a seguir, para a camada de saída. Estas RNAs possuem uma não-linearidade e as principais funções de ativação utilizadas incluem a função logística e a função tangente hiperbólica.

Os MLPs são RNAs diretas (*feedforward*) cujos grafos não possuem ciclos. Os neurônios que recebem sinais de excitação do meio exterior compõem a primeira camada da RNA; uma ou mais camadas intermediárias e a camada de saída. A característica das RNAs em camadas é a organização em grupos, ou seja, cada grupo de neurônios recebe entradas somente do grupo anterior e suas saídas são direcionadas para outro grupo.

A Figura 2.1.3 mostra a arquitetura de uma RNA MLP com uma camada de entrada, uma camada intermediária e uma camada de saída.

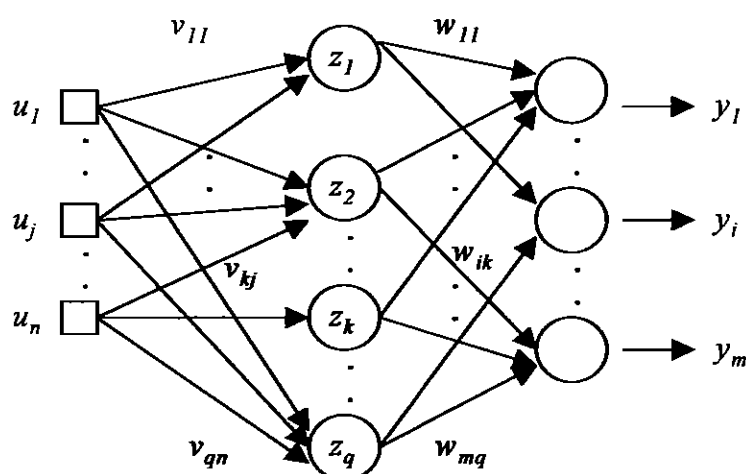


Figura 2.1.3 - Arquitetura de uma RNA *Perceptron* Multicamadas

⁶ Nesta tese é utilizado o termo *retropropagação*.

Observa-se na Figura 2.1.3 que a primeira camada de entrada é passiva, pois apenas recebe os sinais de entrada distribuindo-os para cada neurônio da camada seguinte.

As RNAs MLPs são treinadas de forma supervisionada, utilizando o método baseado no gradiente descendente para obter a direção de busca do mínimo da superfície do erro. Nas RNAs MLPs, o processamento realizado por cada nó é definido pela combinação dos processamentos realizados pelos nós da camada anterior. O principal algoritmo de aprendizagem utilizado para treinar as RNAs MLPs é o de retropropagação⁷, descrito na seção 2.1.4.4.

2.1.4 Formas de Aprendizagem nas RNAs

As formas de aprendizagem são importantes na definição das RNAs, as quais são classificadas em: (1) aprendizagem supervisionada; (2) aprendizagem não supervisionada; (3) aprendizagem híbrida.

Na aprendizagem supervisionada, a saída gerada pelo sistema é comparada com a saída desejada; se forem diferentes, é gerado um sinal de erro e os pesos são reajustados para minimizar o erro gerado. Na aprendizagem não supervisionada, o sinal de entrada é processado e uma saída é gerada, não existindo uma saída desejada, a RNA auto-organiza-se. Já a aprendizagem híbrida combina as duas formas de aprendizagem anteriores.

São vários os algoritmos de aprendizagem, muitos dos quais são baseados numa lei de aprendizagem que se tornou padrão, conhecida como *Lei de Hebb* (Hebb, 1949).

Segue a descrição das formas de aprendizagem hebbiana, regra delta de aprendizagem e regra delta generalizada, as quais forneceram a base para o desenvolvimento desta tese.

· Aprendizagem Hebbiana

Uma das regras de aprendizagem mais antigas é a aprendizagem hebbiana, conhecida como "Lei de Hebb", em razão do neuropsicólogo Donald Hebb, que estudou o comportamento de animais e propôs, em 1949, uma teoria enfatizando que o aprendizado em sistemas nervosos poderia ser reduzido a um processo local, ou seja, que a intensidade das conexões sinápticas pode ser alterada apenas em função dos erros detectáveis localmente (Hebb, 1949).

Uma sinapse hebbiana usa um mecanismo dependente do tempo, altamente local e fortemente interativo para aumentar a eficiência sináptica como uma função da correlação entre as atividades pré-sinápticas (sinal de entrada) e pós-sinápticas (sinal de saída). Nessa forma de aprendizagem vários

⁷ A retropropagação é uma técnica específica que implementa a descida do gradiente no espaço de pesos para uma RNA com múltiplas camadas (Haykin, 2001).

neurônios de saída são ativados simultaneamente e somente um neurônio de saída ficará ativo num determinado instante (Haykin, 2001).

A regra de Hebb, em termos matemáticos, consiste numa regra para alteração do peso sináptico (Haykin, 2001):

$$\Delta w_{ij}^p = \alpha u_j^p y_i^p \quad (2.1.3)$$

onde:

w_{ij}^p = intensidade da conexão do neurônio j para o neurônio i no padrão p ;

Δw_{ij}^p = variação de intensidade da conexão do neurônio j para o neurônio i no padrão p ;

α = taxa de aprendizagem;

u_j^p = entrada do neurônio j para o padrão p ;

y_i^p = saída do neurônio i para o padrão p .

As regras delta e delta generalizada foram inspiradas na regra de Hebb, as quais são utilizadas em alguns algoritmos de aprendizagem, descritas na próxima seção.

· Regra Delta de Aprendizagem

Conhecida como aprendizagem de Widrow-Hoff (Widrow, 1962), consiste numa generalização da regra de Hebb. O princípio básico da regra delta concentra-se em modificar a força das conexões na tentativa de reduzir a diferença (o delta) entre o valor de saída desejado e o valor de saída da RNA. A regra delta é também conhecida como aprendizado *Least Means Square* (LMS). Dessa forma, o erro quadrático total $E(t)$ é minimizado e definido como⁸:

$$E(t) = \sum_p E^p = \frac{1}{2} \sum_p (d_i^p - y_i^p)^2 \quad (2.1.4)$$

onde:

$E(t)$ = erro quadrático total;

p = faixa do conjunto de padrões de entrada;

E^p = erro no padrão p ;

d_i^p = saída desejada i para o padrão p ;

y_i^p = saída atual i para o padrão p .

O procedimento LMS encontra todos os valores de pesos que minimizam a função erro por meio do método chamado de gradiente descendente. A idéia é realizar uma mudança no peso proporcional à negativa da derivada do erro relativo a cada peso para o padrão atual, ou seja:

$$\Delta w_{ij}^p = -\alpha \frac{\partial E^p}{\partial w_{ij}} \quad (2.1.5)$$

⁸ Fórmula adaptada de Haykin (2001).

onde:

α = taxa de aprendizagem da RNA.

O uso do sinal negativo na Equação 2.1.7 refere-se ao gradiente descendente no espaço de pesos, ou seja, à busca pela direção da mudança no peso sináptico de forma que reduza o valor do E^p da RNA. Dessa forma, a derivada é:

$$\frac{\partial E^p}{\partial w_{ij}} = \frac{\partial E^p}{\partial y_i^p} \frac{\partial y_i^p}{\partial w_{ij}} \quad (2.1.6)$$

Em razão das unidades lineares oriundas da Equação $y_i^p = \sum_j w_{ij}u_j + \theta_i$, tem-se:

$$\frac{\partial y_i^p}{\partial w_{ij}} = u_j \quad (2.1.7)$$

$$\frac{\partial E^p}{\partial y_i^p} = \delta_i^p = -(d_i^p - y_i^p) \quad (2.1.8)$$

onde:

θ_i = termo de polarização *bias* da camada de saída i .

Com isso, a regra delta é expressa como:

$$\Delta w_{ij}^p = -\alpha \delta_i^p u_j \quad (2.1.9)$$

onde:

w_{ij}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio i da camada de saída no padrão p ;

Δw_{ij}^p = variação de intensidade da conexão do neurônio j da camada de entrada para o neurônio i da camada de saída no padrão p ;

α = taxa de aprendizagem;

$\delta_i^p = d_i^p - y_i^p$ = diferença entre a saída desejada (d_i^p) e a saída real (y_i^p) do neurônio i para cada padrão p ;

u_j = valor de entrada do neurônio j .

Dessa forma, ocorrerá uma mudança no peso proporcional à negativa do erro relativo a cada peso. A regra delta apresenta uma limitação: é aplicada somente às RNAs que possuem duas camadas de neurônios. Para as RNAs com múltiplas camadas, utiliza-se a regra delta generalizada, descrita na próxima seção.

· *Regra Delta Generalizada*

Conhecida também como algoritmo de retropropagação, a regra delta generalizada constitui uma variação da regra delta. O processo de aprendizagem iterativamente procura por um conjunto de

pesos que minimize a função erro. A regra foi desenvolvida por Werbos (1974) e redescoberta, mais tarde, por Parker (1982), LeCun (1985) e Rumelhart, Hinton e Williams (1986b).

A forma da regra delta generalizada é a mesma da regra delta original, exceto pela adição do termo $\varphi'(net_i^p)$, que constitui a derivada da função de ativação aplicada sobre o valor do potencial de ativação net_i^p . O erro da conexão, a partir do nível mais alto das camadas intermediárias para as unidades de saída, é computado diretamente com a utilização da Equação:

$$\delta_i^p = - (d_i^p - y_i^p) * \varphi'(net_i^p) \quad (2.1.10)$$

onde:

δ_i^p = gradiente local (erro da conexão) do neurônio i para o padrão p ;

$d_i^p - y_i^p$ = diferença entre a saída desejada (d_i^p) e a saída real (y_i^p) do neurônio i para cada padrão p ;

$\varphi'(net_i^p)$ = derivada da função de ativação $\varphi(\cdot)$ avaliada na entrada da RNA net_i^p atingido pelo neurônio i da camada de saída da RNA.

Descrevendo a regra delta generalizada por meio de uma função de ativação logística, tem-se:

$$\varphi(z) = \frac{1}{1 + e^{(-z)}} \quad (2.1.11)$$

cuja derivada é:

$$\frac{d\varphi(z)}{dz} = \varphi'(z) = \varphi(z) * [1 - \varphi(z)] \quad (2.1.12)$$

A função de ativação será uma função diferenciável definida por:

$$y_i^p = \varphi(net_i^p) \quad (2.1.13)$$

na qual

$$net_i^p = \theta_i + \sum_j w_{ij}u_j \quad (2.1.14)$$

onde:

net_i^p = entrada da RNA do neurônio i para o padrão p ;

w_{ij} = intensidade da conexão do neurônio j da camada de entrada para o neurônio i da camada de saída;

u_j = valor de entrada do neurônio j ;

θ_i = representa um valor de polarização (*bias*) da camada de saída i que pode assumir +1 ou -1.

A seguir são derivadas as fórmulas utilizadas no algoritmo de retropropagação a fim de obter a atualização dos pesos da RNA MLP.

Para obter a generalização correta da regra delta tem-se:

$$\Delta w_{ij}^p = -\alpha \frac{\partial E^p}{\partial w_{ij}} \quad (2.1.15)$$

A medida do erro E^p é definida como o erro quadrático total aplicado ao padrão p nas unidades de saída:

$$E(t) = \frac{1}{2} \sum_i (d_i^p - y_i^p)^2 \quad (2.1.16)$$

onde

$$E(t) = \sum_p E^p = \text{somatório do erro quadrático total.}$$

Dessa forma, sua derivada é representada por:

$$\frac{\partial E^p}{\partial w_{ij}} = \frac{\partial E^p}{\partial net_i^p} \frac{\partial net_i^p}{\partial w_{ij}} \quad (2.1.17)$$

A partir da Equação (2.1.14), o segundo fator torna-se:

$$\frac{\partial net_i^p}{\partial w_{ij}} = u_j^p \quad (2.1.18)$$

Quando, então, for definido o erro:

$$\delta_i^p = - \frac{\partial E^p}{\partial net_i^p} \quad (2.1.19)$$

Dessa forma ocorre a atualização da regra delta, que é equivalente à regra delta original, resultando num gradiente descendente na superfície do erro, de forma que o peso se altera em função de:

$$\Delta w_{ij}^p(t) = \alpha \delta_i^p(t) u_j^p(t) \quad (2.1.20)$$

O problema, agora, é saber o que o δ_i^p deve ser para cada unidade i na RNA. O erro é implementado propagando o sinal de erro para trás através da RNA, aplicando a regra da cadeia para escrever a derivada parcial representando o produto de dois fatores, dos quais um reflete mudança no erro como uma função de saída da unidade e o outro, a mudança na saída como uma função de mudança na entrada, apresenta-se como resultado:

$$\delta_i^p = - \frac{\partial E^p}{\partial net_i^p} = - \frac{\partial E^p}{\partial y_i^p} \frac{\partial y_i^p}{\partial net_i^p} \quad (2.1.21)$$

Computando agora para o segundo fator, através da Equação (2.1.13), tem-se:

$$\frac{\partial y_i^p}{\partial net_i^p} = \varphi'(net_i^p) \quad (2.1.22)$$

Para computar o primeiro fator da Equação 2.1.21, são considerados dois casos:

· Primeiro, supõe-se que a unidade i consiste numa unidade de saída da RNA $i = o$. Neste caso segue-se da definição de E^p que:

$$\frac{\partial E^p}{\partial y_o^p} = - (d_o^p - y_o^p) \quad (2.1.23)$$

O resultado obtido é similar ao da regra delta original. Considerando esse resultado às Equações 2.1.22 e 2.1.21, obtém-se:

$$\delta_o^p = (d_o^p - y_o^p) * \varphi'_o(net_o^p) \quad (2.1.24)$$

A Equação 2.1.24 é utilizada para alguma unidade de saída o .

· Segundo, no caso de i não ser uma unidade de saída, mas, sim, uma unidade intermediária k , a medida do erro é escrita como uma função das entradas da RNA a partir da camada intermediária para a camada de saída. Então, $E^p = E^p (net_1^p, net_2^p, \dots, net_h^p, \dots)$ e, por meio da regra da cadeia, escreve-se:

$$\frac{\partial E^p}{\partial y_i^p} = \sum_h \frac{\partial E^p}{\partial net_h^p} \frac{\partial net_h^p}{\partial y_i^p} \quad (2.1.25)$$

$$= \sum_h \frac{\partial E^p}{\partial net_h^p} \frac{\partial}{\partial y_i^p} \sum_k w_{hk} y_k^p \quad (2.1.26)$$

$$= \sum_h \frac{\partial E^p}{\partial net_h^p} w_{hi} \quad (2.1.27)$$

$$= - \sum_h \delta_h^p w_{hi} \quad (2.1.28)$$

A Equação 2.1.21 passa a ser:

$$\delta_i^p = \varphi' (net_i^p) \sum_h \delta_h^p w_{hi} \quad (2.1.29)$$

As Equações 2.1.24 e 2.1.29 fornecem o procedimento recursivo para computar os erros (δ 's) para todas as unidades da RNA através da atualização dos pesos e de acordo com a Equação 2.1.15.

Esse procedimento constitui numa regra delta generalizada para uma RNA *feedforward*.

Na seqüência, descreve-se o algoritmo de retropropagação utilizado para treinar as RNAs MLPs⁹.

2.1.5 Algoritmo de Retropropagação

Segue a nomenclatura de códigos, adaptada de Fausett (1994):

w_j^p = valor de entrada j para o padrão p

d_i^p = valor desejado de saída i para o padrão p

δ_i^p = erro da unidade de saída i para o padrão p

δ_k^p = retropropagação do erro a partir da camada de saída i até a camada intermediária k para o padrão p

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p

α = taxa de aprendizagem

θ_k^p = *bias* da camada intermediária k para o padrão p

⁹ O algoritmo de *retropropagação* é descrito nesta tese por ser utilizado na avaliação dos modelos de aprendizagem investigados.

Z_k^p = valor de saída da camada intermediária k para o padrão p

$\varphi(\cdot)$ = função de ativação

$\varphi'(\cdot)$ = derivada da função de ativação

=> a entrada da RNA para Z_k^p é dada por: $net_k^p = \theta_k + \sum_j (v_{kj}u_j)$

=> o sinal de saída (ativação) Z_k^p é dado por z_k^p :

$$\Rightarrow z_k^p = \varphi(net_k^p)$$

θ_i^p = bias da camada de saída i para o padrão p

$\varphi(net_i^p)$ = função de ativação avaliada na entrada da RNA net_i^p atingida pelo neurônio i da camada de saída da RNA para o padrão p

$\varphi(net_k^p)$ = função de ativação avaliada na entrada da RNA net_k^p atingida pelo neurônio k da camada intermediária da RNA para o padrão p

Y_i^p = valor de saída da camada de saída i para o padrão p

=> a entrada da RNA para Y_i^p é dada por: $net_i^p = \theta_i + \sum_k (w_{ik}z_k)$

=> o sinal de saída (ativação) de Y_i^p é denotado por y_i^p :

$$\Rightarrow y_i^p = \varphi(net_i^p)$$

A seguir faz-se uma descrição do algoritmo de retropropagação, dividido em três fases: inicialização, *feedforward* e *backward*.

I – Fase de Inicialização:

Etapa_1: repetir as etapas 2 a 9 até a condição de parada;

Etapa_2: para cada par do treinamento (d_i, u_j) , executar as etapas 3 a 8;

II – Fase Feedforward:

Etapa_3: cada unidade de entrada $(u_j, j = 1, \dots, n)$ recebe o sinal transmitindo-o para as camadas escondidas;

Etapa_4: cada unidade escondida $(Z_k, k = 1, \dots, q)$ soma os pesos dos sinais de entrada:

$net_k^p = \theta_k + \sum_j (v_{kj}u_j)$, após, é aplicada a função de ativação para computar a saída do sinal

$z_k^p = \varphi(net_k^p)$ a seguir o sinal é propagado para a camada de saída

Etapa_5: cada unidade de saída $(Y_i, i = 1, \dots, m)$ realiza a soma dos pesos dos sinais de entrada:

$net_i^p = \theta_i + \sum_k (w_{ik}z_k^p)$, após aplica a função de ativação para computar o sinal de saída

$$y_i^p = \varphi(net_i^p)$$

III – Fase Backward:

Etapa_6: cada unidade de saída ($Y_i, i = 1, \dots, m$) recebe a saída desejada correspondente ao padrão de treinamento e computa o seu erro;

O objetivo é minimizar o erro usando o método do gradiente descendente:

$\delta_i^p = (d_i^p - y_i^p) * \varphi'(net_i^p)$, a seguir são calculados os seus pesos para, depois, atualizar w_{ik}

$\Delta w_{ik}^p = \alpha * \delta_i^p * z_k^p$, após é calculado o *bias*, o qual será utilizado para atualizar θ_i

$\Delta \theta_i^p = \alpha * \delta_i^p$, a seguir envia δ_i^p para as unidades da camada abaixo

Etapa_7: cada unidade intermediária ($Z_k, k = 1, \dots, q$) realiza a soma das entradas Δ ;

$\Delta \delta_k^p = \sum_i (\delta_i^p * \Delta w_{ik}^p)$, a seguir a soma das entradas Δ é multiplicada pela derivada de sua função para calcular a formação de erro

$\delta_k^p = \Delta \delta_k^p * \varphi'(net_k^p)$, realiza o cálculo dos seus pesos, os quais serão atualizados na próxima etapa

$\Delta v_{kj}^p = \alpha * \delta_k^p * u_j^p$, após executa o cálculo do *bias* para depois atualizar θ_k

$\Delta \theta_k^p = \alpha * \delta_k^p$

Etapa_8: fase de atualização dos pesos e do *bias*;

Cada unidade de saída ($Y_i, i = 1, \dots, m$) atualiza os pesos w_{ik} e o *bias* θ_i

$$w_{ik}^{p+1} = w_{ik}^p + \Delta w_{ik}^p$$

$$\theta_i^{p+1} = \theta_i^p + \Delta \theta_i^p$$

Cada unidade escondida ($Z_k, k = 1, \dots, q$) atualiza os pesos v_{kj} e o *bias* θ_k

$$v_{kj}^{p+1} = v_{kj}^p + \Delta v_{kj}^p$$

$$\theta_k^{p+1} = \theta_k^p + \Delta \theta_k^p$$

Etapa_9: testar a condição de parada.

2.1.6 Considerações Finais

Atualmente, observa-se crescente interesse pela pesquisa em RNAs, as quais apresentam soluções eficientes na resolução de problemas de aprendizagem e na classificação de bases de conhecimento.

A partir do referencial desta seção, é proposta nesta tese uma RNA híbrida baseada na combinação de RNAs e LD. O modelo de RNA escolhido é o MLP, uma vez que ele tem sido aplicado e difundido entre as arquiteturas de RNAs mais conhecidas. A opção por uma RNA híbrida objetiva investigar a utilização dos métodos difusos para intensificar a capacidade de aprendizagem da RNA. Nesta RNA híbrida são trabalhados os Conjuntos Difusos (CDs) e as variáveis lingüísticas, as quais

são utilizadas no pré-processamento das informações de entrada juntamente com a implementação de neurônios difusos.

A próxima seção está relacionada com a LD, descrevendo suas principais características.

2.2 Lógica Difusa

A LD fornece um método matemático para representar a imprecisão associada com os sistemas cognitivos humanos, tais como o pensamento e o raciocínio. A LD procura simular a forma do pensamento humano utilizando palavras ou expressões no lugar de valores numéricos. Através dela, pode-se realizar operações com palavras, onde os CDs são os valores das palavras (Braga, Barreto e Machado, 1995).

Esta seção fornece uma breve introdução aos principais conceitos relacionados com a LD.

2.2.1 Considerações Iniciais

A LD tem como objetivo a formalização de um raciocínio aproximado, ao invés de exato - utilizado na representação de informações cujos limites não são precisos. Ou seja, constitui-se de uma ferramenta capaz de capturar informações imprecisas, descritas em uma linguagem natural e transformá-las para um formato numérico.

Os fundamentos da LD foram desenvolvidos por Zadeh (1965), baseados na Teoria dos Conjuntos Tradicionais e na Lógica Clássica¹⁰. L.A. Zadeh e R. Bellman verificaram a necessidade de se criar uma teoria que trabalhasse com a imprecisão em sistemas dinâmicos. A LD permite a utilização de um intervalo de valores de pertinência entre os limites 0 e 1, distinguindo-se, dessa forma, da Teoria Clássica dos Conjuntos. Utilizando CDs pode-se trabalhar não apenas com os valores 0 e 1, mas com qualquer valor entre esses limites, ou seja, no intervalo $[0, 1]$. A função característica de um conjunto *crisp* (abrupto), isto é, o conjunto da Teoria Clássica dos Conjuntos, possui valores 1 ou 0 para cada indivíduo do Conjunto do Universo U , definindo, assim, se o indivíduo é membro ou não do conjunto.

¹⁰ Conhecido como o *Father of Fuzzy Logic*, introduziu o conceito da utilização de conjuntos cujos limites não são precisos, ou seja, são utilizados graus de pertinência para definir se um objeto pertence ou não (com diferentes intensidades) a um determinado conjunto.

2.2.2 Conjuntos Difusos

Um CD A é uma classe de conjuntos de objetos com graus de pertinência contínuos, a qual designa a cada objeto x um grau de pertinência ao Conjunto A , denotado por $\mu_A(x)$, que está entre o intervalo 0 e 1 (Zadeh, 1965). A partir deste conceito, tem-se:

$$A = \{(U, \mu_A(x)) \mid x \in U\} \quad (2.2.1)$$

onde:

A = é um conjunto difuso;

$\mu_A(x)$ = grau de pertinência ou função de pertinência do objeto x ao conjunto difuso A ;

U = universo de discurso.

De acordo com a Equação 2.2.1, o grau de pertinência associa cada objeto x , pertencente a U , um número real $\mu_A(x)$ no intervalo $[0,1]$, representado pelo grau de pertinência de que o elemento x pertença ao conjunto difuso A .

Os CDs expressam o significado dos valores lingüísticos relacionados com uma variável lingüística. Como exemplo, a variável lingüística de entrada da RNA "desmaio"¹¹ pode estar associada a um conjunto de valores lingüísticos, como "às_vezes", "frequentemente" ou "sempre".

As variáveis lingüísticas, definidas como CDs, representam uma faixa específica de valores, ou seja, quando se define, por exemplo, que o paciente "desmaiou", isso não significa uma resposta "afirmativa" ou "negativa", mas graus de pertinência.

A representação dos CDs depende da dimensão do universo de discurso definido, ou seja, quando o universo for discreto e composto por um número pequeno de objetos, a maneira adotada é através da representação analítica, descrita de acordo com Rezende et al. (2003):

$$\mu_A(x) = \{1, 0/0; 0, 8/1; 0, 6/2; 0, 4/3; 0, 2/4; 0, 0/5\} \quad (2.2.2)$$

Na Equação 2.2.2, cada objeto do Conjunto A é representado na forma $\mu_A(x_i)$, onde o primeiro termo representa o grau de pertinência de x_i ao Conjunto A e o segundo termo identifica o objeto x_i .

Quando o universo de discurso for contínuo, a forma de representação faz uso do diagrama de Hassi-Euler (H-E), como mostra a Figura 2.2.1.

¹¹ A variável lingüística "desmaio" representa um sintoma utilizado no diagnóstico dos eventos epiléticos.

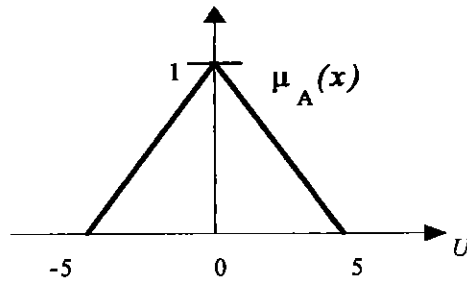


Figura 2.2.1 - Representação da função de pertinência (adaptado de Rezende et al. (2003))

A Figura 2.2.1 exemplifica o Conjunto A no universo de discurso U associado a um grau de pertinência. Dessa forma, um objeto poderá pertencer (com diferentes intensidades) ou não a um CD. O grau de pertinência pode variar no intervalo de 0 a 1, ou seja, da completa exclusão até a total pertinência. Na Figura 2.2.1 tem-se que:

- $\forall (x) \quad | \quad x < -5, \quad x \text{ não pertence ao Conjunto } A;$
- $\forall (x) \quad | \quad -5 \leq x < 0, \quad x \text{ pertence cada vez mais ao Conjunto } A;$
- para* $x = 0, \quad x \text{ pertence completamente ao Conjunto } A;$
- $\forall (x) \quad | \quad 0 \leq x < 5, \quad x \text{ pertence cada vez menos ao Conjunto } A;$
- $\forall (x) \quad | \quad x \geq 5, \quad x \text{ não pertence ao Conjunto } A.$

Dessa forma, o grau de pertinência num CD é obtido pela generalização da função característica, chamada função de pertinência, definida por:

$$\mu_A(x): x \rightarrow [0, 1] \tag{2.2.3}$$

A função de pertinência mapeia x no domínio dos números reais definidos no intervalo 0 a 1, inclusive.

O grau de pertinência é um número real definido por Klir e Yuan (1995):

$$0 \leq \mu_A(x) \leq 1 \tag{2.2.4}$$

onde:

0 = não pertinência do objeto x ao Conjunto A ;

1 = total pertinência do objeto x ao Conjunto A .

Ou seja, um CD é aquele no qual a pertinência de seus elementos não ocorre de forma absoluta (pertence ou não pertence). O objeto pode pertencer ou não a um conjunto com um determinado grau de pertinência.

2.2.3 Operações com Conjuntos Difusos

Uma classe geral de operadores de intersecção para CDs é definida pelas "normas triangulares" ou *t-normas*, e uma classe de operadores de união é definida, analogicamente, pelas *t-conormas* ou *S-normas*. As *t-normas* são aplicadas na LD para modelar o conectivo lógico *E*; similarmente, as *t-conormas* modelam o conectivo lógico *OU*.

As operações aritméticas difusas padrões para as *t-normas* e as *t-conormas*, considerando os Conjuntos *A* e *B* (Figura 2.2.2), são descritas de acordo com Klir e Yuan (1995):

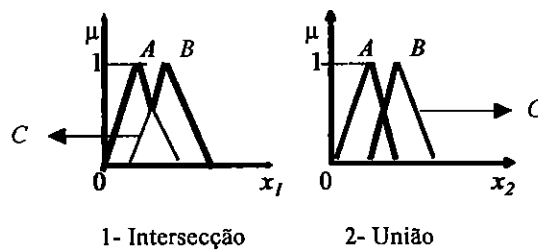


Figura 2.2.2 - Operações com conjuntos difusos

onde:

(1) intersecção, $\mu_C(x) = \min \{ \mu_A(x_1), \mu_B(x_1) \}$;

(2) união, $\mu_C(x) = \max \{ \mu_A(x_2), \mu_B(x_2) \}$.

A operação intersecção (*t-norma*) entre CDs apresenta as propriedades:

- Comutatividade: $T(a, b) = T(b, a)$;
- Associatividade: $T(a, T(b, c)) = T(T(a, b), c)$;
- Monotonicidade: se $a \leq b$ e $c \leq d$, então $T(a, c) \leq T(b, d)$;
- Coerência nos contornos: $T(a, 1) = a$ e $T(a, 0) = 0$.

A operação de união (*t-conormas*) entre CDs apresenta as propriedades:

- Comutatividade: $S(a, b) = S(b, a)$;
- Associatividade: $S(a, S(b, c)) = S(S(a, b), c)$;
- Monotonicidade: se $a \leq b$ e $c \leq d$, então $S(a, c) \leq S(b, d)$;
- Coerência nos contornos: $S(a, 1) = 1$ e $S(a, 0) = a$.

Além das operações aritméticas padrões *max/min*, nesta tese foram exploradas diferentes operações aritméticas difusas *t-normas* e *t-conormas* na fase de avaliação dos modelos de aprendizagem, de acordo com Klir e Yuan (1995):

Produto Algébrico:

$$\mu_C(x) = (\mu_A(x) \cdot \mu_B(x)) \quad (2.2.5)$$

Soma Algébrica:

$$\mu_C(x) = (\mu_A(x) + \mu_B(x)) - (\mu_A(x) \cdot \mu_B(x)) \quad (2.2.6)$$

Produto de Hamacher:

$$\mu_C(x) = \frac{\mu_A(x) \cdot \mu_B(x)}{(\mu_A(x) + \mu_B(x)) - (\mu_A(x) \cdot \mu_B(x))} \quad (2.2.7)$$

Soma de Hamacher:

$$\mu_C(x) = \frac{\mu_A(x) + \mu_B(x) - 2 \mu_A(x) \cdot \mu_B(x)}{1 - (\mu_A(x) \cdot \mu_B(x))} \quad (2.2.8)$$

Produto de Einstein:

$$\mu_C(x) = \frac{\mu_A(x) \cdot \mu_B(x)}{2 - [((\mu_A(x) + \mu_B(x))) - ((\mu_A(x) \cdot \mu_B(x)))]} \quad (2.2.9)$$

Soma de Einstein:

$$\mu_C(x) = \frac{\mu_A(x) + \mu_B(x)}{1 + (\mu_A(x) \cdot \mu_B(x))} \quad (2.2.10)$$

Exemplificando as operações com CDs, os graus de pertinência são combinados e trabalhados utilizando as operações aritméticas difusas (Equações 2.2.5 à 2.2.10), ou seja, quando for utilizada a operação aritmética padrão *max*, representada pela *união/OU*, o resultado será o maior grau de pertinência entre eles, isto é, $\max(x_1, x_2, \dots, x_n)$: se $\mu_A = 0,2$; $\mu_B = 0,5$, neste caso tem-se $\mu_A \cup \mu_B = \max(0,2; 0,5) = 0,5$. Quando for utilizada a operação aritmética padrão *min*, representada pela *intersecção/E*, a resposta será o menor grau de pertinência, ou seja, $\min(x_1, x_2, \dots, x_n)$: neste caso tem-se $\mu_A \cap \mu_B = \min(0,2; 0,5) = 0,2$.

2.2.4 Base de Regras

As regras que compõem o sistema podem ser fornecidas pelo especialista ou ser extraídas de dados numéricos. A base de regras representa a maneira como o sistema será gerenciado, apresentam-se na forma *SE-ENTÃO*, onde a parte *SE* está relacionada aos termos antecedentes e a parte *ENTÃO*, aos termos conseqüentes.

Os graus de pertinência para um conjunto de regras poderão ser colocados tanto na parte antecedente como na parte conseqüente. Exemplificando, o grau de pertinência na parte antecedente é definido com a seguinte regra:

· *SE* dor na garganta "leve" (0, 2) *OU* "estável" (0, 4) *OU* dificuldade de engolir (0, 6) *ENTÃO* *Conclusão*. Aplicando a operação aritmética difusa padrão *max*, tem-se:

$$R1 = \max (0, 2; 0, 4; 0, 6) = 0, 6 \tag{2.2.11}$$

Quando o grau de pertinência for implementado nas duas partes (antecedente e conseqüente), a regra apresenta a seguinte descrição:

· *SE* dor na garganta "leve" (0, 2) *OU* "estável" (0, 4) *OU* dificuldade de engolir (0, 6) *ENTÃO* *Conclusão* (0, 4) . O grau de pertinência obtido será:

$$R2 = \max ((\max(0, 2; 0, 4; 0, 6)); (0, 4)) = 0, 6 \tag{2.2.12}$$

2.2.5 Sistemas de Inferência Difusos

Os Sistemas de Inferência Difusos (SIDs) são ferramentas computacionais baseados nos conceitos da teoria dos CDs, regras (*SE-ENTÃO*) e raciocínio difuso (Jang, Sun e Mizutani, 1997).

Na literatura são encontrados os seguintes SIDs: (1) *Mamdani* (1974); (2) *Takagi Sugeno Kang* (1985, 1988); (3) *Tsukamoto* (1979). A proposta metodológica da tese faz uso do SID *Mamdani*, uma vez que as operações aritméticas difusas padrões *max/min*, implementadas com a ferramenta NEFCLASS, utilizam o SID *Mamdani*. Posteriormente, essas operações aritméticas difusas padrões são substituídas por outras operações aritméticas difusas, conforme descrito na seção 2.2.3. O SID *Mamdani* é a única abordagem aqui apresentada. As operações aritméticas difusas investigadas são comparadas com as operações aritméticas difusas padrões *max/min*. No SID *Mamdani* os antecedentes e conseqüentes de cada regra são CDs representando variáveis lingüísticas de entrada e de saída do sistema.

A estrutura básica do SID *Mamdani* é exemplificada na Figura 2.2.3.

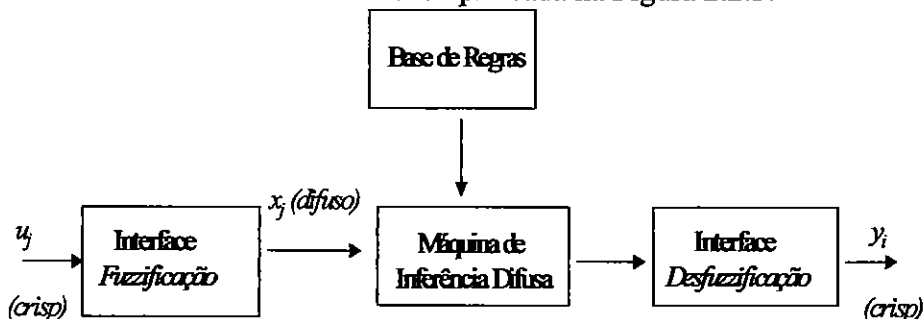


Figura 2.2.3 - SID *Mamdani* (adaptado de Gorzalczany (2002))

Observa-se na Figura 2.2.3 que o SID é composto por quatro componentes conceituais:

- (1) interface de *fuzzificação*;
- (2) máquina de inferência difusa;
- (3) base de regras e
- (4) interface de *desfuzzificação*.

A interface de *fuzzificação* calcula o grau de pertinência a partir de um valor de entrada *crisp* (u_j) em um ou mais CDs, resultando num valor de saída difuso (x_j). A máquina de inferência executa o raciocínio difuso pela combinação das regras (*SE-ENTÃO*), de forma a encontrar o conseqüente e o seu grau de pertinência. A base de regras é composta por uma seleção das regras e contém funções de pertinência definidas tanto para os termos antecedentes como para os conseqüentes. A interface de *desfuzzificação* transforma um conjunto difuso de saída e seu grau de pertinência num valor *crisp* (y_i), através de algum método de *desfuzzificação*¹². O desenvolvimento dessa interface justifica-se apenas quando se deseja obter um valor numérico de saída.

As regras utilizadas no processamento de inferências no modelo *Mamdani* são chamadas de inferência *max/min* e utilizam as operações intersecção (*t-normas*) e/ou as operações união (*t-conormas*) entre CDs.

As Equações 2.2.13 e 2.2.14 descrevem uma regra genérica exemplificando o método de inferência *max/min* (adaptado de Rezende et al., 2003):

$$\text{SE } x_1 \text{ igual } A_1 \text{ E } x_2 \text{ igual } A_2 \text{ E } \dots \text{ E } x_n \text{ igual } A_n \quad (2.2.13)$$

$$\text{ENTÃO } y_1 = B_1 \text{ OU } y_2 = B_i \text{ OU } \dots \text{ OU } y_m \text{ igual } B_m \quad (2.2.14)$$

onde:

x_j = valor de entrada *fuzzificado* j do sistema;

$A_1, \dots, A_j, \dots, A_n$ = CDs dos termos antecedentes;

y_1 e y_2 = valores de saída;

$B_1, \dots, B_i, \dots, B_m$ = CDs dos termos conseqüentes.

A Figura 2.2.4 ilustra o SID *Mamdani*, composto por duas regras difusas, realizando o processo de inferência difusa *max/min*.

¹² Os métodos de *desfuzzificação* não são descritos nesta tese porque estão fora do escopo dessa pesquisa.

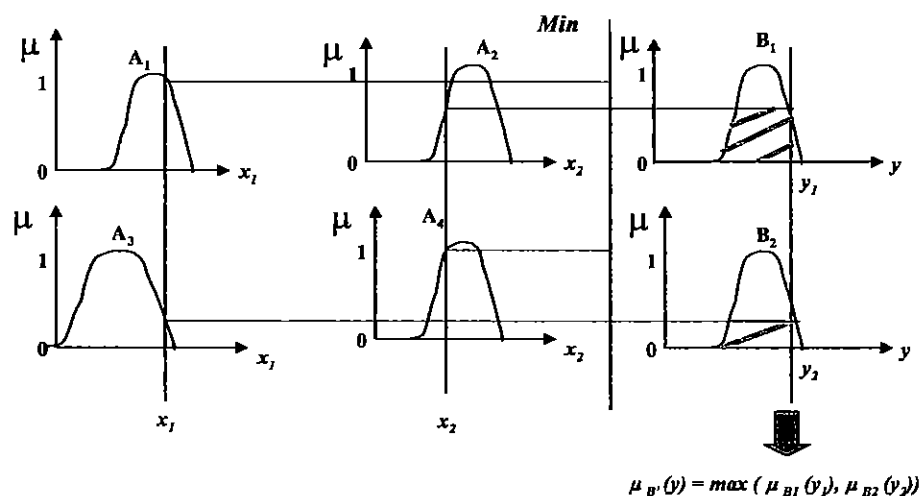


Figura 2.2.4 - SID *Mamdani* (adaptado de Melin e Castillo (2005))

Na Figura 2.2.4, os CDs consequentes B_1 e B_2 representam o resultado da operação *min* entre os CDs antecedentes A_1, A_2 e A_3 e A_4 , respectivamente. O CD consequente B' representa o resultado da operação *max* entre os CDs consequentes B_1 e B_2 .

O SID apresenta as seguintes regras:

· SE x_1 igual A_1 E x_2 igual A_2

ENTÃO y_1 igual B_1

· SE x_1 igual A_3 E x_2 igual A_4

ENTÃO y_2 igual B_2

inferindo as regras,

$$\mu_{B_1}(y_1) = \min(\mu_{A_1}(x_1), \mu_{A_2}(x_2))$$

$$\mu_{B_2}(y_2) = \min(\mu_{A_3}(x_1), \mu_{A_4}(x_2))$$

concluindo,

$$\mu_{B'}(y) = \max(\mu_{B_1}(y_1), \mu_{B_2}(y_2)).$$

2.2.6 Considerações Finais

Nesta seção foi descrito o SID *Mamdani*, pois os modelos de aprendizagem explorados são comparados com as operações aritméticas difusas padrões *max/min* por meio do SND que utiliza a arquitetura NEFCLASS.

As abordagens envolvendo RNAs e a LD apresentam mecanismos de inferência diferentes e trabalham com informações imprecisas. Um dos grandes desafios das pesquisas é superar algumas deficiências relacionadas com as RNAs e a LD, por exemplo, utilizando algoritmos de aprendizagem que possam incorporar operações aritméticas difusas numa RNA para ambas as abordagens, neurais e difusas, compartilharem das vantagens presentes em cada uma.

A partir do referencial acima, nesta tese foram implementadas e avaliadas as operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein OU(E), juntamente com as operações aritméticas difusas padrões *max/min* através da implementação dos neurônios difusos propostos por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994), aplicadas numa RNA híbrida.

Esse tema, explorando o hibridismo entre RNAs e a LD, por ser o foco principal desta tese, é descrito em detalhes na próxima seção.

2.3 Sistemas Neuro-Difusos

Esta seção descreve os principais tópicos relacionados com os modelos de aprendizagem utilizados nos SNDs.

2.3.1 Considerações Iniciais

Historicamente, as primeiras pesquisas relacionadas com os SNDs foram propostas por Lee e Lee (1974), publicadas na década de 1974. Porém, foi no trabalho pioneiro dos pesquisadores H. Takagi e I. Hayashi, na década de 1980, que a teoria se consolidou (Rutkowska, 2002).

Quando são utilizados métodos difusos para intensificar a capacidade de aprendizagem das RNAs e esses métodos são combinados numa arquitetura híbrida, têm-se os SNDs, os quais podem ser implementados utilizando-se diferentes algoritmos de aprendizagem.

No desenvolvimento de um SND são utilizados métodos de aprendizagem que implementam neurônios difusos, envolvendo as etapas de modelagem das conexões sinápticas, as quais incorporam conceitos difusos numa RNA e o desenvolvimento do algoritmo de aprendizagem representado pelo método de ajuste dos pesos sinápticos (Fullér, 2000).

2.3.2 Tipos de SNDs

Na literatura são encontradas diferentes definições envolvendo o trabalho conjunto entre RNAs e a LD. Uma descrição destas definições inclui os temas (Nauck e Kruse, 1994_{a,b}):

- *Sistemas Neuro-Difusos Concorrentes*: uma RNA e um SID trabalham juntos na mesma tarefa, todavia um não influenciando o outro. Normalmente, a RNA executa o pré-processamento das entradas ou o pós-processamento das saídas, a partir do sistema difuso.

- *Sistemas Neuro-Difusos Cooperativos*: neste modelo uma RNA é utilizada para determinar alguns parâmetros (regras, pesos ou conjuntos difusos) num SID. Após a fase de aprendizagem, o SID trabalha sozinho, ou seja, sem a utilização da RNA.

- *Sistemas Híbridos Neuro-Difusos*: uma RNA e um SID são combinados numa arquitetura híbrida. O sistema interpreta uma RNA através de parâmetros difusos ou um SID implementado de forma paralela e distribuída. Alguns exemplos deste sistema são encontrados nos sistemas ANFIS (Jang, 1993), NEFCLASS (Nauck e Kruse, 1995) e NEFCON (Nauck e Kruse, 1994_b).

Nesta tese foi utilizado um SND híbrido, pois a RNA implementada trabalha de forma paralela com o SID, o qual interpreta a RNA utilizando parâmetros difusos (regras e CDs).

São descritos dois modelos de SNDs propostos por Fullér (2000):

A Figura 2.3.1 exemplifica o primeiro SND, onde o bloco da interface difusa fornece um vetor de entrada para uma RNA multicamada.

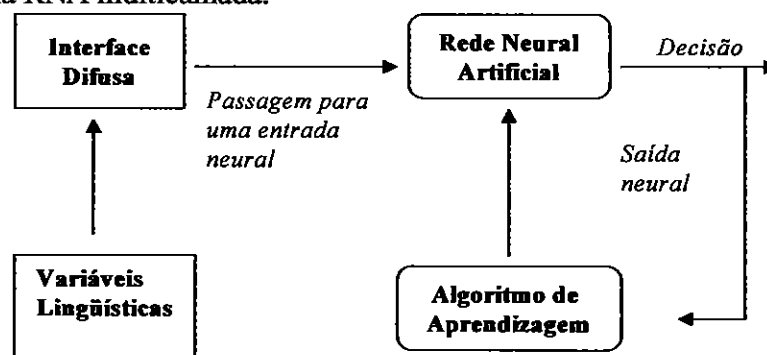


Figura 2.3.1 - Primeiro SND (adaptado de Fullér (2000))

Após as variáveis lingüísticas serem *fuzziificadas* (interface difusa) inicia-se o processo de treinamento da RNA, utilizando um algoritmo de aprendizagem.

No segundo SND (Figura 2.3.2) uma RNA realiza o mecanismo de inferência difuso.

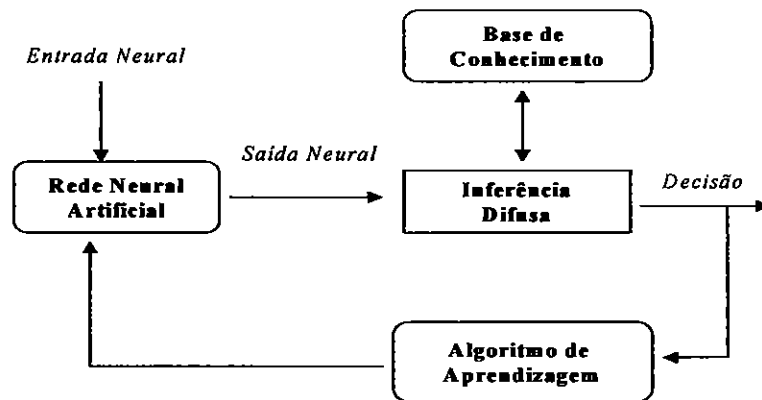


Figura 2.3.2 - Segundo SND (adaptado de Fullér (2000))

Na Figura 2.3.2, a RNA é utilizada no ajuste das funções de pertinência do sistema difuso.

Osório e Vieira (2007) definem os SNDs em três grupos:

- 1) Sistemas que integram regras difusas nas RNAs, processo chamado de "compilação de regras em RNAs";
- 2) Sistemas que realizam a extração das regras difusas a partir das RNAs, os quais extraem automaticamente os conhecimentos adquiridos por uma RNA, sob a forma de regras da lógica difusa;
- 3) Sistemas que implementam neurônios com processamento difuso, conhecidos como "neurônios difusos", os quais são transformados em elementos do processamento realizando operações aritméticas similares às operações utilizadas na LD.

A Figura 2.3.3 apresenta um modelo de SND sugerido por Kartalopoulos (1996).

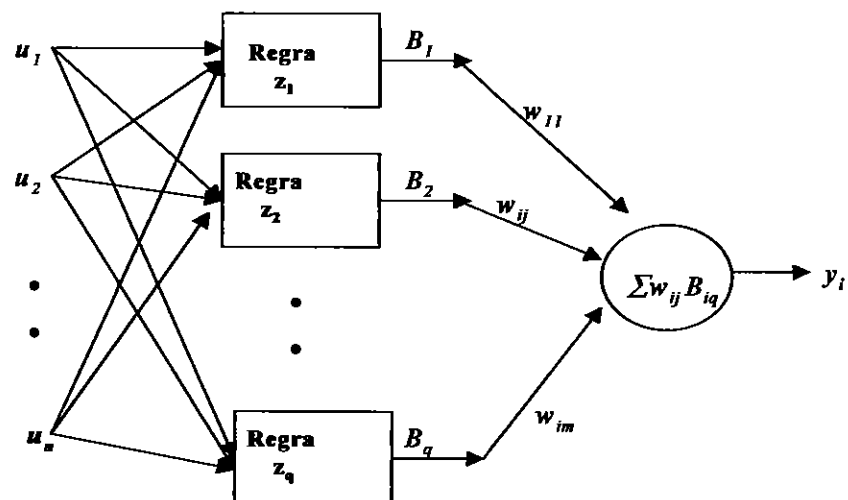


Figura 2.3.3 - Modelo de SND (adaptado de Kartalopoulos (1996))

Na Figura 2.3.3 são apresentados os seguintes elementos:

(a) a primeira camada representa as entradas do sistema (u_1, u_2, \dots, u_n) ; (b) a segunda camada representa a camada de inferência utilizando as regras (*SE-ENTÃO*) definidas como (z_1, z_2, \dots, z_q) :

$$SE \ x_1 \text{ igual } A_1 \ E, \dots, \ E \ x_j \text{ igual } A_j \ E, \dots, \ E \ x_n \text{ igual } A_n \quad (2.3.1)$$

$$ENTÃO \ y_1 = B_1 \ E \ y_2 = B_2 \ E, \dots, \ E \ y_m \text{ igual } B_q \quad (2.3.2)$$

Uma saída y_i^p é gerada a partir da soma dos pesos dos valores de entrada:

$$y_i^p = \sum_j (w_{ij} B_{iq}) \quad (2.3.3)$$

onde:

x_j = valor de entrada *fuzzificado* j ;

y_i^p = saída i para o padrão p .

A escolha por um ou outro sistema depende da aplicação proposta.

Nesta tese é utilizado o termo SND incluindo o hibridismo presente nas RNAs e na LD.

Na próxima seção são descritos os mecanismos envolvidos no projeto de SNDs.

2.3.3 Projeto de Sistemas Neuro-Difusos

O principal elemento no processamento de uma RNA é o neurônio artificial. O sinal flui através dos neurônios de entrada $(u_1, \dots, u_j, \dots, u_n)$ interagindo com os seus pesos $(w_{11}, \dots, w_{ij}, \dots, w_{im})$ para produzir um produto $p_i^p = w_{ij}u_j$, $j = 1, \dots, n$. Essa informação de entrada p_i^p é somada para produzir a entrada do neurônio (Fullér, 2000):

$$net_i^p = p_1^p + \dots + p_i^p + \dots + p_m^p = w_{11}u_1 + \dots + w_{ij}u_j + \dots + w_{im}u_n \quad (2.3.4)$$

A seguir, é utilizada uma função de ativação $\varphi(.)$ como, por exemplo, a função logística:

$$\varphi(net_i^p) = \frac{1}{1 + e^{(-x)}} \quad (2.3.5)$$

para computar a saída da RNA:

$$y_i^p = \varphi(net_i^p) = \varphi(p_1^p + \dots + p_i^p + \dots + p_m^p) \quad (2.3.6)$$

Ao serem utilizadas outras operações aritméticas difusas *t-normas* ou *t-conormas*, no lugar do produto $(w_{ij}u_j)$, tem-se o que se chama de "RNA híbrida". Essas modificações levam ao projeto de SNDs baseados em operações aritméticas difusas, e o seu principal elemento de processamento é

o neurônio difuso. As entradas da RNA após a fuzzificação são representadas por meio de graus de pertinência $(x_1, \dots, x_j, \dots, x_n)$ e os pesos podem estar no intervalo $[-1,1]$ (Fullér, 2000).

Na próxima seção faz-se uma definição dos modelos de neurônios difusos.

2.3.4 Neurônios Difusos

Na literatura são encontrados vários modelos de neurônios difusos, os quais envolvem a manipulação dos operadores lógicos $OU(E)$ e das operações aritméticas difusas t -normas e t -conormas, respectivamente. A implementação dos neurônios difusos difere na forma de manipular as operações aritméticas difusas.

A seguir, a descrição dos modelos de neurônios difusos proposta por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994), os quais são utilizados nesta tese.

· *Modelo Hirota e Pedrycz (1992, 1994)*

Hirota e Pedrycz (1992, 1994) propõem os neurônios difusos E/OU , os quais combinam os sinais de entrada e os pesos utilizando as operações aritméticas difusas t -normas e t -conormas, respectivamente¹³. Após, o resultado é transferido para a saída, utilizando uma função de ativação.

A Figura 2.3.4 exemplifica o neurônio difuso E ¹⁴.

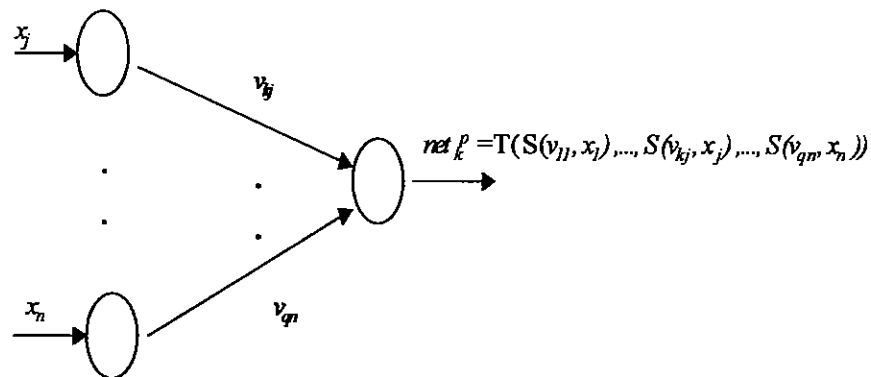


Figura 2.3.4 - Neurônio difuso E (adaptado de Hirota e Pedrycz (1992, 1994))

onde:

x_j = valor de entrada fuzzificado j ;

x_j e v_{kj} são combinados por uma operação t -conorma S para produzir p_k^j :

$$p_k^j = S(v_{kj}^p, x_j^p), j = 1, \dots, n \tag{2.3.7}$$

¹³ Alguns autores chamam essas operações aritméticas difusas de "Função de Agregação" (Hirota e Pedrycz, 1992, 1994).

¹⁴ Os exemplos definem os neurônios a partir da camada de entrada j para a camada intermediária k .

onde:

p_k^p = resultado da operação *t-conorma* S aplicada no neurônio k da camada intermediária para o padrão p .

A entrada da informação p_k^p é agregada por uma operação *t-norma* T para produzir a saída net_k^p :

$$net_k^p = T(S(p_{11}^p), \dots, S(p_{kj}^p), \dots, S(p_{qn}^p)) \quad (2.3.8)$$

A Figura 2.3.5 exemplifica o neurônio difuso *OU*:

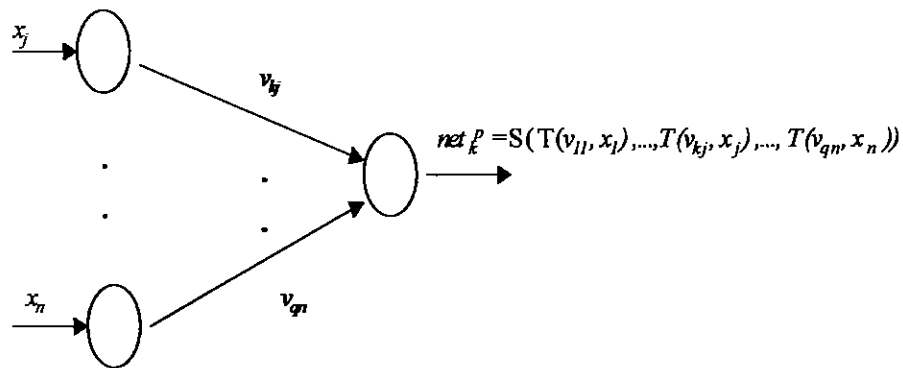


Figura 2.3.5 - Neurônio difuso *OU* (adaptado de Hirota e Pedrycz (1992, 1994))

onde:

x_j e v_{kj} são combinados por uma operação *t-norma* T para produzir p_k^p :

$$p_k^p = T(v_{kj}, x_j), \quad j = 1, \dots, n \quad (2.3.9)$$

A entrada da informação p_k^p é agregada por uma operação *t-conorma* S para produzir a saída net_k^p :

$$net_k^p = S(T(p_{11}^p), \dots, T(p_{kj}^p), \dots, T(p_{qn}^p)) \quad (2.3.10)$$

onde:

p_k^p = resultado da operação aritmética difusa *t-norma* T aplicada no neurônio k da camada intermediária para o padrão p ;

net_k^p = entrada da RNA atingido pelo neurônio k da camada intermediária da RNA para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p ;

x_j^p = valor de entrada *fuzzificado* j para o padrão p .

A seguir, o modelo de neurônio difuso proposto por Kwan-Cai (1994).

· Modelo Kwan e Cai (1994)

Kwan e Cai (1994) propõem os neurônios difusos E/OU , utilizando, além das operações aritméticas difusas t -normas e t -conormas, o produto das entradas com os seus pesos ($v_{kj} x_j$).

A Figura 2.3.6 exemplifica o neurônio difuso E^{15} .

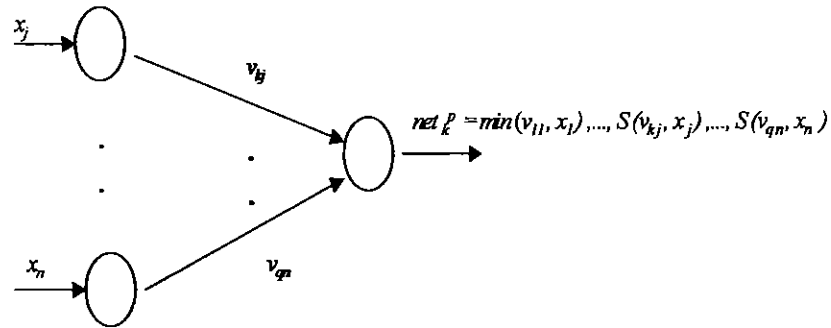


Figura 2.3.6 - Neurônio difuso E (adaptado de Kwan e Cai (1994))

Na Figura 2.3.6, o neurônio difuso E referente à camada intermediária k , apresenta o sinal de entrada *fuzzificado* x_j interagindo com o peso sináptico v_{kj} para produzir o produto p_k^p :

$$p_k^p = (v_{kj}x_j), j = 1, \dots, n \tag{2.3.11}$$

onde:

p_k^p = resultado da operação *produto* aplicada no neurônio k da camada intermediária para o padrão p .

A entrada da informação p_k^p é agregada por uma operação aritmética difusa t -norma T (min) para computar a saída net_k^p :

$$net_k^p = \min \{ (p_{11}^p), \dots, (p_{kj}^p), \dots, (p_{qn}^p) \} = T \{ v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n \}$$

A saída do neurônio z_k^p é computada por:

$$z_k^p = (\varphi(net_k^p - \theta_k^p)) = (\varphi(T \{ v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n \} - \theta_k^p)) \tag{2.3.12}$$

onde:

z_k^p = valor de saída da camada intermediária k para o padrão p ;

$\varphi(\cdot)$ = função de ativação logística;

net_k^p = entrada da RNA atingida pelo neurônio k da camada intermediária da RNA para o padrão p ;

15 Os exemplos definem os neurônios a partir da camada de entrada j para a camada intermediária k .

$\theta_k^p = threshold$ de ativação da camada intermediária k para o padrão p ;

$v_{kj}^p = intensidade$ da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p ;

$x_j^p = valor$ de entrada *fuzzificado* j para o padrão p .

A Figura 2.3.7 exemplifica o neurônio difuso *OU*:

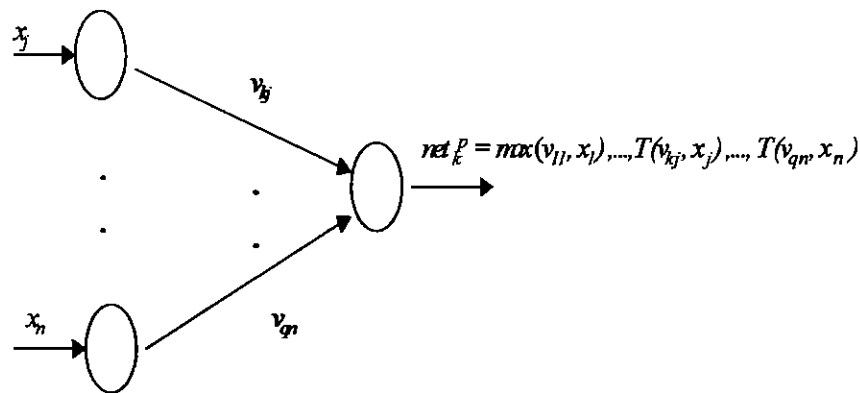


Figura 2.3.7 - Neurônio difuso *OU* (adaptado de Kwan e Cai (1994))

Na Figura 2.3.7, no neurônio difuso *OU* o sinal de entrada *fuzzificado* x_j , interage com o peso sináptico v_{kj} para produzir o produto p_k^p :

$$p_k^p = v_{kj} x_j, j = 1, \dots, n \tag{2.3.13}$$

A entrada da informação p_k^p é agregada por uma operação aritmética difusa *t-conorma* S (*max*) para produzir a saída net_k^p :

$$net_k^p = max \{ (p_{11}^p), \dots, (p_{kj}^p), \dots, (p_{qn}^p) \} = S \{ v_{11} x_1, \dots, v_{kj} x_j, \dots, v_{qn} x_n \} \tag{2.3.14}$$

A saída do neurônio z_k^p é computada por:

$$z_k^p = (\varphi(net_k^p - \theta_k^p)) = (\varphi(S \{ v_{11} x_1, \dots, v_{kj} x_j, \dots, v_{qn} x_n \} - \theta_k^p)) \tag{2.3.15}$$

Outros modelos de neurônios difusos encontrados na literatura são os de Yager (1988, 1991, 1992), Yamakawa (1989) e Kartalopoulos (1996). A principal diferença entre os modelos está na forma de manipular as operações aritméticas difusas *t-normas* e *t-conormas*.

Na próxima seção são descritas algumas arquiteturas de SNDs encontradas na literatura.

2.3.5 Arquiteturas de Sistemas Neuro-Difusos

Esta seção apresenta algumas arquiteturas de SNDs que serviram de base para o processo de avaliação dos modelos de aprendizagem propostos nesta tese.

Os SNDs descritos integram os formalismos de representação de conhecimento através de seus métodos de inferência, operações aritméticas difusas e operadores lógicos *E/OU*.

· *Arquitetura ANFIS (Jang, 1993)*

Uma das primeiras arquiteturas de SNDs referenciada na literatura foi o *Adaptative Network based Fuzzy Inference Systems (ANFIS)*, introduzido por Jang, Sun e Mizutani (1993,1997).

A representação do ANFIS segue o SID TSK sugerido por Takagi, Sugeno e Kang (1985, 1988) de primeira ordem (Jang, Sun e Mizutani, 1997):

$$SE x_1 igual A_1 E x_2 igual A_3, ENTÃO f_1 = p_1 x_1 + q_1 x_2 + r_1 \quad (2.3.16)$$

$$SE x_1 igual A_2 E x_2 igual A_4, ENTÃO f_2 = p_2 x_1 + q_2 x_2 + r_2 \quad (2.3.17)$$

onde:

x_1 e x_2 = valores *fuzzificados* de entrada da RNA;

A_1, A_2, A_3 e A_4 = CDs dos antecedentes;

f_1 e f_2 = resultados dos polinômios para as entradas *fuzzificadas* x_1 e x_2 ;

p_1, p_2, q_1, q_2, r_1 e r_2 = termos de um polinômio de primeira ordem.

A arquitetura ANFIS utiliza uma RNA *feedforward* com cinco camadas, algoritmo de retropropagação e aprendizagem supervisionada (Figura 2.3.8).

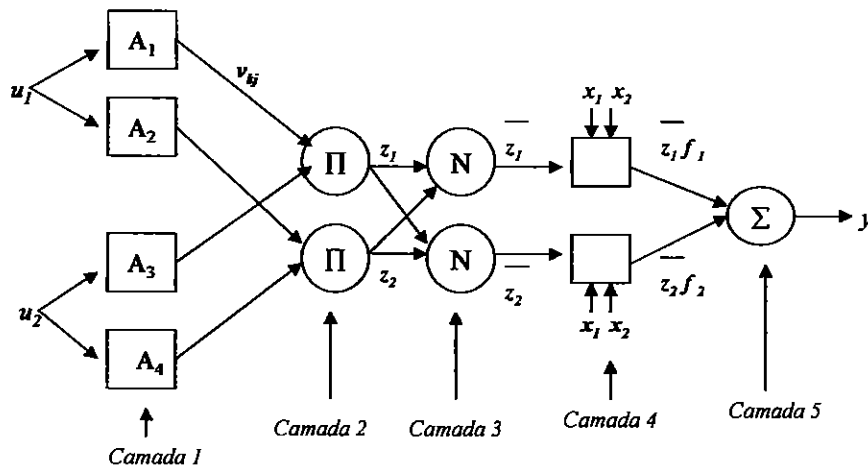


Figura 2.3.8 - Arquitetura ANFIS (adaptada de Jang, Sun e Mizutani (1997))

onde:

- os nós quadrados são adaptativos, ou seja, apresentam parâmetros modificáveis;
- os nós circulares são fixos, ou seja, não apresentam parâmetros ajustáveis.

Na Figura 2.3.8 a arquitetura ANFIS contém duas unidades de entrada (u_1 e u_2). As camadas, denotadas por $Camada_1, \dots, Camada_5$, apresentam as seguintes funções:

- $Camada_1$: o objetivo desta camada é encontrar o grau de pertinência dos nós de entrada através dos CDs, no intervalo $[0,1]$.

$$(x_1) = \mu_{A_j}(u_1) \text{ para } j = 1, 2 \text{ e} \quad (2.3.18)$$

$$(x_2) = \mu_{A_j}(u_2) \text{ para } j = 3, 4 \quad (2.3.19)$$

onde:

u_1 e u_2 = valores de entrada da RNA;

- $Camada_2$: o objetivo desta camada é calcular a força de disparo das regras implementadas no sistema, onde a saída representa uma operação *t-norma*. O operador E das Equações 2.3.17 e 2.3.18 representa a operação aritmética difusa *t-norma* produto algébrico¹⁶:

$$(z_1) = \mu_{A_j}(x_1) \cdot \mu_{A_j}(x_2) \text{ para } j = 1 \text{ e } j = 3 \quad (2.3.20)$$

$$(z_2) = \mu_{A_j}(x_1) \cdot \mu_{A_j}(x_2) \text{ para } j = 2 \text{ e } j = 4 \quad (2.3.21)$$

Na Figura 2.3.8 o sinal Π representa o produtório.

- $Camada_3$: as saídas das unidades desta camada representam valores normalizados a partir da ativação das regras:

$$(\bar{z}_k) = \frac{z_k}{\sum_{k=1}^2 z_k} \text{ para } k = 1, 2 \quad (2.3.22)$$

Na Figura 2.3.8 o sinal N significa normalização:

- $Camada_4$: os nós nesta camada são nós que calculam o conseqüente de uma regra (saída):

$$(\bar{z}_k f_k) = \bar{z}_k (p_k x_1 + q_k x_2 + r_k) \text{ para } k = 1, 2 \quad (2.3.23)$$

onde:

\bar{z}_k = valores normalizados de ativação das regras a partir da $Camada_3$;

$p_k x_1 + q_k x_2 + r_k$ = termos de um polinômio de primeira ordem representando os parâmetros conseqüentes.

¹⁶ Outras operações aritméticas difusas *t-normas* podem ser utilizadas nessa camada (Jang, 1993).

· *Camada₅*: esta camada realiza a saída total através do somatório de todos os sinais que estão chegando a partir da *Camada₄*:

$$y = \sum (\bar{z}_k f_k) \text{ para } k = 1, 2 \quad (2.3.24)$$

Vários SNDs encontrados na literatura são baseados na arquitetura ANFIS, entre eles os SNDs propostos por Aguiar, Melin e Castillo (2003), Lim et al., (2005) e Yu, Lai e Lee (2006).

A arquitetura ANFIS encontra-se disponível na URL: <<http://www-2.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/fuzzy/systems/anfis/>>.

· *Arquitetura Mitra e Pal (1992)*

Esta arquitetura integra uma RNA difusa, baseada no MLP, utilizando o algoritmo de retropropagação na classificação de padrões e geração de regras. A Figura 2.3.9 descreve os valores de entrada da RNA ($u_1, \dots, u_j, \dots, u_n$). Após a fuzzificação as entradas *fuzzificadas* da RNA ($x_1, \dots, x_j, \dots, x_n$) apresentam graus de pertinências considerando o intervalo [0,1]. As saídas da RNA ($y_1, \dots, y_i, \dots, y_m$) representam as classes de padrões associadas com os seus graus de pertinência (Mitra e Pal, 1992, 1994, 1995, 1997).

A RNA é treinada a partir de exemplos, generalizando-os na fase de teste com exemplos desconhecidos e que não foram treinados. Durante a fase de treinamento é utilizada a aprendizagem supervisionada para determinar com precisão o valor do grau de pertinência de saída, que se encontra na faixa de [0, 1].

Os pesos sinápticos são atualizados de acordo com a Equação:

$$\Delta w_{ik}^{p+1} = n(\Delta w_{ik}^{p-1}) + \alpha * \delta_i^p * z_k^p \quad (2.3.25)$$

onde:

n = constante 1;

Δw_{ik}^{p-1} = variação da intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão $p - 1$;

α = taxa de aprendizagem;

δ_i^p = erro da unidade de saída i para o padrão p ;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

O algoritmo de retropropagação convencional foi modificado para incorporar o operador lógico *E* (*t-norma*) e o operador lógico *OU* (*t-conorma*). Durante a fase de treinamento, a taxa de aprendizagem decresce gradualmente até a RNA convergir para um erro mínimo.

A Figura 2.3.9 exemplifica a estrutura da RNA contendo três camadas com múltiplas saídas.

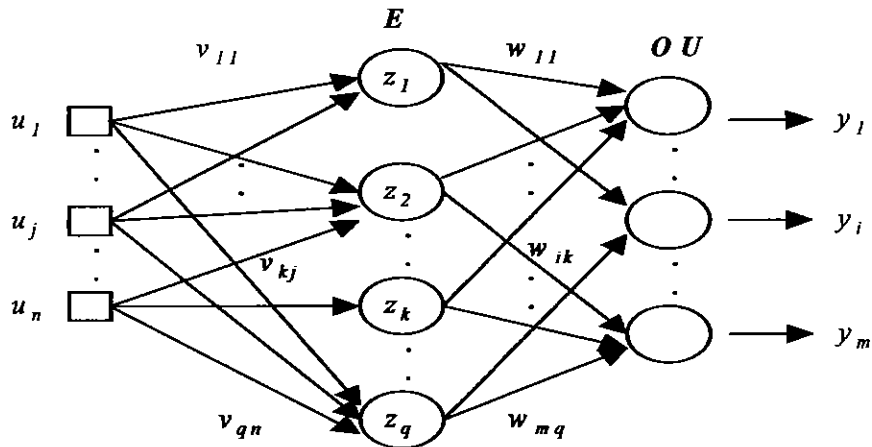


Figura 2.3.9 - Operadores lógicos E/OU (adaptado de Mitra e Pal (1992))

De acordo com a Figura 2.3.9, a camada intermediária k consiste de q neurônios E , e a camada de saída i , de m neurônios OU . Após o treinamento da RNA, a mesma é utilizada na geração de regras difusas, generalizando-a na fase de teste. Na fase de explicação, o usuário solicita ao sistema por que foi inferida uma determinada conclusão. O sistema, então, responde através de uma regra *SE-ENTÃO*, gerada pelo sistema de inferência a partir dos pesos das conexões sinápticas.

· *Arquitetura FuNet-I (Halgamuge e Glesner, 1994)*

Halgamuge e Glesner (1994) propõem uma arquitetura baseada no MLP, aprendizagem supervisionada, conhecida como *FuNe-I*. A arquitetura *FuNe-I* é composta por cinco camadas: (1) a primeira camada representa os valores de entrada ($u_1, \dots, u_j, \dots, u_n$); (2) a segunda camada representa a fuzzificação dos valores de entrada com seus respectivos CDs A_j , resultando nas entradas *fuzzificadas* da RNA ($x_1, \dots, x_j, \dots, x_n$), utilizando função de ativação logística na obtenção dos graus de pertinência; (3) a terceira camada contém unidades especializadas utilizadas para representar combinações de CDs; (4) a quarta camada representa as regras utilizando as operações aritméticas difusas *Softmax/Softmin* e (5) uma quinta camada integra as unidades de saída da RNA, as quais implementam suas entradas utilizando uma soma de pesos e suas funções de ativação, por uma função logística (Figura 2.3.10).

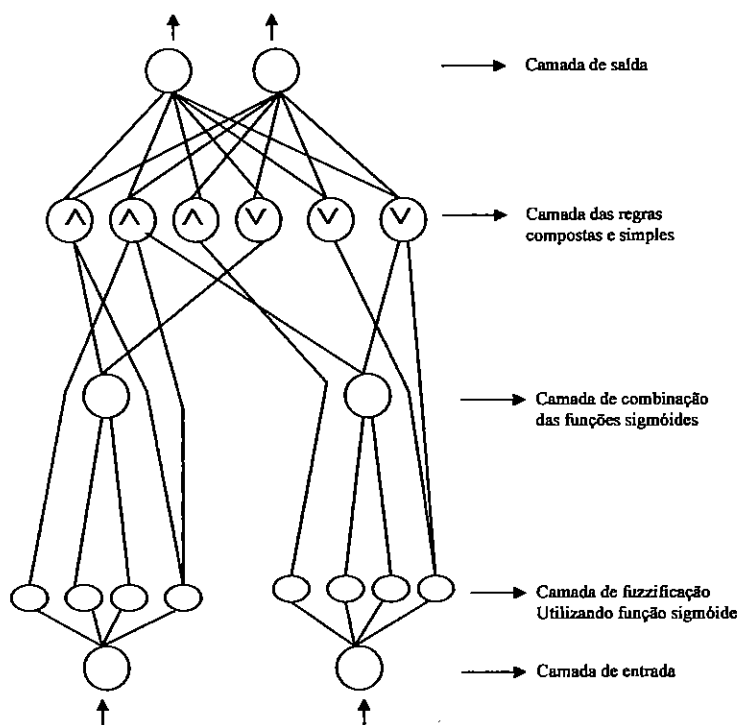


Figura 2.3.10 - SND FuNe-I (adaptado de Halgamuge e Glesner (1994))

Observa-se na Figura 2.3.10 que algumas unidades da segunda camada apresentam conexões para a quarta camada, que não são conectadas à terceira camada. Para cada unidade de entrada existem quatro unidades na segunda camada, onde uma unidade representa o CD "pequeno"; e uma outra unidade, o CD "grande". As outras duas unidades, junto com uma unidade na terceira camada, representam o CD "médio". Estas funções de pertinência que não são localizadas nos limites do domínio "médio" são representadas por duas funções logísticas.

A arquitetura FuNe-I utiliza dois tipos de regras: regras compostas, cujos antecedentes podem ser conjuntivos ou disjuntivos, e regras simples, com somente uma variável no termo antecedente.

O processo de aprendizagem é baseado no algoritmo de retropropagação modificado com os operadores lógicos *E/OU*. Os autores utilizam as operações aritméticas difusas *Softmax/Softmin* em substituição às operações aritméticas difusas padrões *max/min*, por serem deriváveis e também por apresentarem facilidade de implementação, descritas por meio das seguintes Equações (Berenji e Khedkar, 1992):

$$Softmin(x_1, \dots, x_j, \dots, x_n) = \frac{\sum_{j=1}^n x_j * e^{-cx_j}}{\sum_{j=1}^n e^{-cx_j}} \tag{2.3.26}$$

e

$$Softmax(x_1, \dots, x_j, \dots, x_n) = \frac{\sum_{j=1}^n x_j * e^{cx_j}}{\sum_{j=1}^n e^{cx_j}} \tag{2.3.27}$$

onde:

e = função exponencial;

c = variável que pode ser incrementada ao infinito;

x_j = valor de entrada *fuzzificado* j .

· *Arquitetura Gupta e Rao (1994_{ab})*

Esta arquitetura descreve um SND baseado nas operações aritméticas difusas *t-normas* e *t-conormas* na construção do neurônio difuso utilizando o algoritmo de aprendizagem de retropropagação (Gupta e Rao, 1994_{ab}).

O SND possui uma topologia de RNA direta com múltiplas camadas, apresentando neurônios difusos *OU* na camada intermediária e na camada de saída da RNA (Figura 2.3.11).

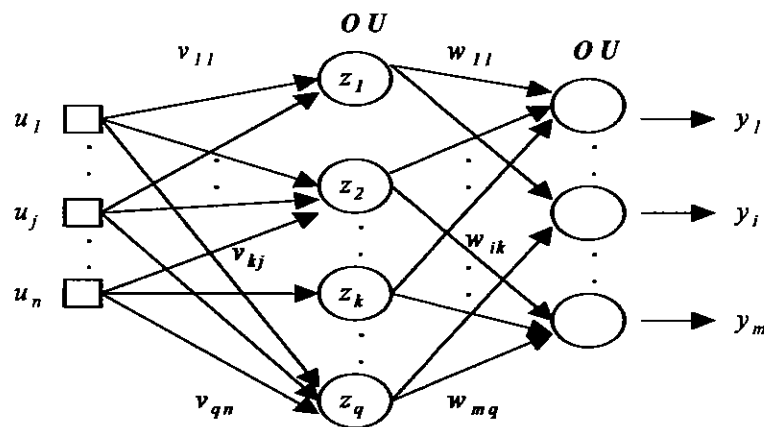


Figura 2.3.11 - SND com neurônios OU (adaptado de Brasil (1999))

O ambiente computacional utilizado na construção SND inicia com o desenvolvimento de um neurônio difuso. Os CDs representam os termos lingüísticos "pequeno", "médio" e "grande".

O processo de aprendizagem compreende três etapas: (1) desenvolvimento de um modelo neural difuso; (2) modelagem das conexões sinápticas; (3) desenvolvimento de um algoritmo de aprendizagem para ajustar os pesos sinápticos.

A regra de aprendizagem para uma RNA supervisionada utiliza o algoritmo de retropropagação; para uma RNA não supervisionada, é utilizada a regra de aprendizado de Hebb.

A entrada neural difusa é descrita em termos de graus de pertinência, conforme a Equação:

$$x_j^p = (\theta_k, x_1, \dots, x_j, \dots, x_n) \in [0, 1]$$

onde:

x_j^p = valor de entrada *fuzzificado* j para o padrão p ;

θ_k = sinal de polarização *bias* da camada intermediária k .

As entradas *fuzzificadas* da RNA $(x_1, \dots, x_j, \dots, x_n)$, juntamente com seus pesos, foram substituídas pelo operador *t-norma* T e a operação somatório, pelo operador *t-conorma* S .

$$T : [0, 1] \times [0, 1] \longrightarrow [0, 1] \quad (2.3.28)$$

$$net_i^p = [x_1 \ E \ x_2] \triangleq [x_1 T \ x_2] = T [x_1, x_2] \quad (2.3.29)$$

$$S : [0, 1] \times [0, 1] \longrightarrow [0, 1] \quad (2.3.30)$$

$$net_i^p = [x_1 \ OU \ x_2] \triangleq [x_1 S \ x_2] = S [x_1, x_2] \quad (2.3.31)$$

As entradas *fuzzificadas* e os pesos sinápticos estão no intervalo $[0, 1]$, definindo os sinais unipolares compostos por interações excitatórias positivas, ou no intervalo $[-1, 1]$, definindo os sinais bipolares. As funções de ativação utilizadas são a logística e a tangente hiperbólica.

Para a atualização dos pesos do neurônio difuso *OU* tem-se:

$$net_i^p = S[w_{ik}^p \ T \ z_k^p] \in [0, 1] \quad (2.3.32)$$

onde:

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p ;

z_k^p = valor de saída da camada intermediária k para o padrão p .

O erro na camada de saída é definido como:

$$E(t) = \frac{1}{2} \sum_i (d_i^p - y_i^p)^2 \in [-1, 1] \quad (2.3.33)$$

onde:

$E(t)$ = erro quadrático total ;

d_i^p = saída desejada i para o padrão p ;

y_i^p = saída real i para o padrão p .

Adaptando-se a função erro acima no algoritmo de retropropagação, tem-se:

$$\Delta w_{ik}^{p+1} = w_{ik}^p + \Delta w_{ik}^p \quad (2.3.34)$$

onde:

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p ;

Δw_{ik}^p = mudança no peso sináptico w_{ik}^p ;

Dessa forma, para a atualização das conexões sinápticas dos neurônios OU , tem-se:

$$\Delta w_{ik}^{p+1} = w_{ik}^p S \Delta w_{ik}^p \quad (2.3.35)$$

$$\Delta w_{ik}^{p+1} = S [w_{ik}^p, \Delta w_{ik}^p] \quad (2.3.36)$$

$$\Delta w_{ik}^{p+1} = \max [w_{ik}^p, \Delta w_{ik}^p] \quad (2.3.37)$$

Para a atualização das conexões sinápticas envolvendo neurônios E , tem-se:

$$\Delta v_{kj}^{p+1} = v_{kj}^p T E(t) \quad (2.3.38)$$

$$\Delta v_{kj}^{p+1} = T [v_{kj}^p, E(t)] \quad (2.3.39)$$

$$\Delta v_{kj}^{p+1} = \min [v_{kj}^p, E(t)] \quad (2.3.40)$$

onde:

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p ;

Δv_{kj}^p = mudança de peso sináptico v_{kj}^p ;

$E(t)$ = erro quadrático total.

· *Arquitetura Pedrycz* (1991, 1993)

Pedrycz (1991, 1993) introduziu novos modelos de neurônios com aprendizagem supervisionada. A arquitetura utiliza o algoritmo de aprendizagem de retropropagação modificado com a inclusão de algoritmos genéticos na tarefa de otimizar o aprendizado da RNA. O modelo neural do neurônio

incorpora diferentes funções triangulares utilizando as operações aritméticas difusas *t-normas* e *t-conormas*.

O autor sugere dois tipos de arquiteturas de SNDs. Inicialmente, os SNDs propostos implementavam as operações aritméticas padrões *max/min* com o objetivo de compartilhar recursos envolvendo as RNAs e a LD. Posteriormente, além de analisar os resultados das operações aritméticas difusas padrões *max/min*, o autor sugere a utilização de neurônios de referência, baseados em predicados lógicos em relação a um ponto de referência (Pedrycz e Rocha, 1993, Pedrycz, 2004).

Na seqüência descrevem-se as pesquisas envolvendo as operações aritméticas difusas *t-normas* e *t-conormas*.

A primeira arquitetura proposta utiliza neurônios difusos *E/OU* com as operações aritméticas padrões *max/min*. O neurônio difuso *E* representa a camada intermediária e o neurônio difuso *OU*, a camada de saída, como mostra a Figura 2.3.12.

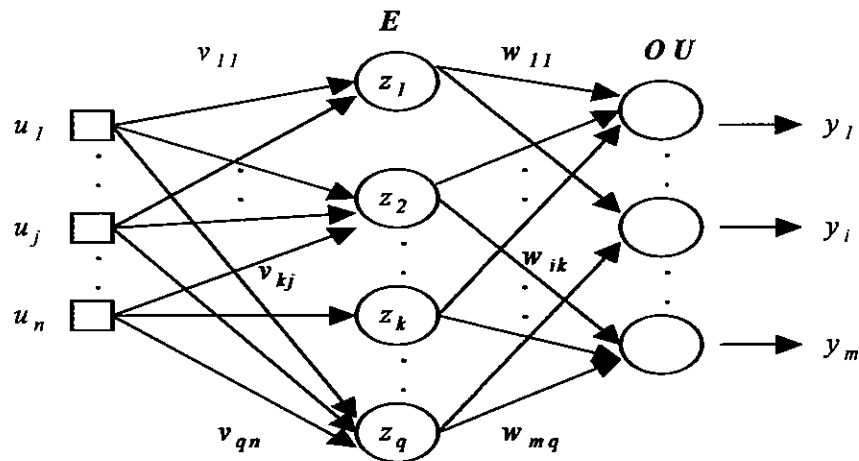


Figura 2.3.12 - Inferência difusa E/OU (adaptado de Pedrycz (1993))

Na segunda arquitetura, a camada intermediária é composta por neurônios difusos *E/OU* e a camada de saída contém somente neurônios difusos *OU*, juntamente com as operações aritméticas difusas padrões *max/min*. Ambas as arquiteturas utilizam funções triangulares na modelação dos seus CDs.

Os neurônios difusos *E* são descritos como:

Os sinais de entrada *fuzzificados* $(x_1, \dots, x_j, \dots, x_n)$ são combinados com os seus respectivos pesos sinápticos $(v_{11}, \dots, v_{kj}, \dots, v_{qn})$, utilizando o conectivo lógico *OU* para cada par (x_j^p, v_{kj}^p) . A seguir, esses resultados são agregados através do conectivo lógico *E*, produzindo a saída da camada

intermediária da RNA net_k^p . Essa transformação é descrita como:

$$net_k^p = \prod_{j=1}^n (x_j^p S v_{kj}^p) \quad (2.3.41)$$

ou

$$net_k^p = \min[\max(x_j^p, v_{kj}^p)] \quad (2.3.42)$$

onde:

net_k^p = saída da RNA atingida pelo neurônio k da camada intermediária para o padrão p ;

x_j^p = valor de entrada *fuzziificado* j para o padrão p ;

v_{kj}^p = intensidade da conexão sináptica do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p .

· Os neurônios difusos *OU* são descritos como:

Este neurônio constitui-se numa dupla estrutura, ou seja, os sinais de entrada são inicialmente combinados com os seus pesos sinápticos ($w_1, \dots, w_i, \dots, w_m$), utilizando o conectivo lógico *E* para cada par $(net_k^p, w_{ik}^p)^{17}$. A seguir, os resultados são agregados através do conectivo lógico *OU*, produzindo a saída da RNA net_i^p . Essa transformação é descrita de acordo com:

$$net_i^p = \sum_{k=1}^q (net_k^p T w_{ik}^p) \quad (2.3.43)$$

ou

$$net_i^p = \max[\min(net_k^p, w_{ik}^p)] \quad (2.3.44)$$

onde:

net_i^p = saída da RNA atingida pelo neurônio i da camada de saída para o padrão p ;

w_{ik}^p = intensidade da conexão sináptica do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p .

As funções são especificadas através das operações aritméticas difusas *t-norma* (*min*) e *t-conorma* (*max*), respectivamente.

· *Arquitetura NEFCLASS* (Nauck e Kruse, 1996)

A arquitetura *Neuro-Fuzzy Classification* (*NEFCLASS*), baseada no MLP, foi introduzida por Nauck e Kruse (1995, 1996) para resolver problemas de classificação de padrões. A arquitetura é treinada com um conjunto de padrões, onde cada elemento pertence a uma classe distinta. Após o processo de aprendizagem, a arquitetura NEFCLASS é utilizada novamente para classificação de dados previamente desconhecidos.

¹⁷ Neste exemplo são considerados como entradas os sinais da camada intermediária k da rede.

Nesta tese optou-se pela implementação da arquitetura NEFCLASS na fase de avaliação dos resultados, uma vez que o NEFCLASS utiliza as operações aritméticas difusas padrões (*max/min*). Dessa forma, seus resultados são comparados com as demais operações aritméticas difusas *t-normas* e *t-conormas* implementadas no SisND/EPs.

As regras de classificação são implementadas de acordo com:

$$\text{SE } x_1 \text{ igual } A_1 \text{ E } \dots, x_j \text{ igual } A_j \text{ E } \dots, \text{ E } x_n \text{ igual } A_n \text{ ENTÃO } y_i \quad (2.3.45)$$

onde:

x_j = entradas *fuzziificadas* da RNA;

$A_1, \dots, A_j, \dots, A_n$ = valores lingüísticos "às_vezes", "freqüentemente" e "sempre";

y_i = classes de saída da RNA.

A Figura 2.3.13 apresenta a arquitetura NEFCLASS.

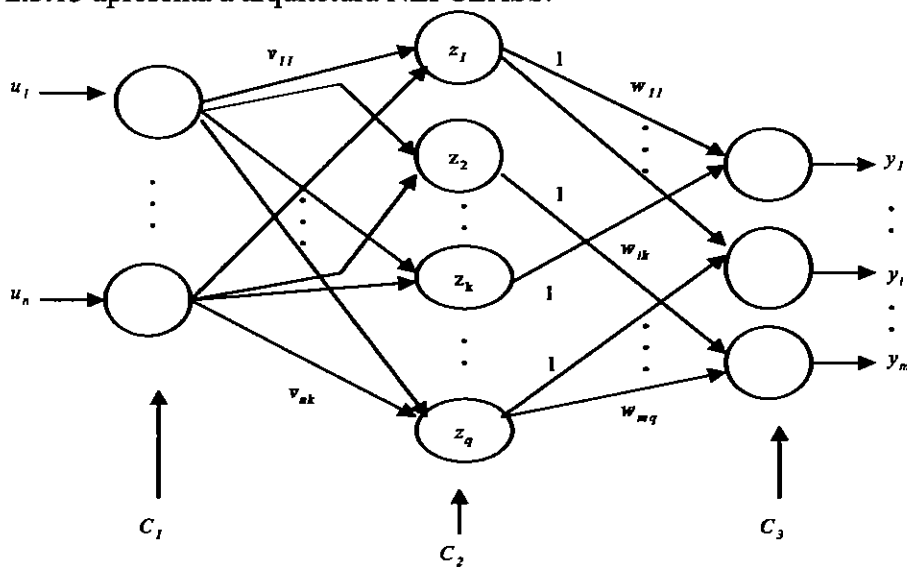


Figura 2.3.13 - Arquitetura NEFCLASS (adaptada de Nauck e Kruse (1995, 1996))

O sistema possui uma arquitetura de RNA *feedforward* com três camadas, denominadas de C_1, C_2 e C_3 . A camada C_1 representa os valores *fuzziificados* de entrada ($x_1, \dots, x_j, \dots, x_n$), descrevendo as características dos padrões a serem classificados.

Na camada C_2 são implementadas as regras de classificação utilizando as operações aritméticas difusas padrões *max/min*. O objetivo desta camada é calcular o grau de ativação de uma regra para um determinado padrão de entrada e contém um nó para cada regra *SE-ENTÃO*. A camada C_3 representa as unidades de saída, uma para cada classe. Os pesos das conexões (w_{ik}) a partir das

camadas C_2 e C_3 são fixados em 1. Cada unidade na camada C_3 calcula seu valor de entrada como a soma dos pesos das saídas dessas unidades na camada C_2 .

Na arquitetura NEFCLASS os CDs A_j são representados por uma função de pertinência triangular:

$$\mu_{A_j}(x_j) = \frac{u_j - a}{b - a} \quad \text{se } u_j \in [a, b] \quad (2.3.46)$$

ou

$$\mu_{A_j}(x_j) = \frac{c - u_j}{c - b} \quad \text{se } u_j \in [b, c] \quad (2.3.47)$$

ou

$$\mu_{A_j}(x_j) = 0 \quad \text{se outro valor} \quad (2.3.48)$$

onde:

$\mu_{A_j}(x_j)$ = grau de pertinência do objeto x_j ao Conjunto A_j ;

a, b, c = parâmetros dos CDs.

Em virtude de as operações aritméticas difusas padrões *max/min* serem funções não-deriváveis, o algoritmo de aprendizagem do NEFCLASS não utiliza o método do gradiente descendente, e, sim, apenas a função sinal do gradiente (*sgn*), de acordo com as Equações 2.3.51, 2.3.52 e 2.3.53 (Nauck, 1994):

· cada unidade de saída recebe a saída desejada, correspondente ao padrão de treinamento e computa o seu erro δ_i^p :

$$E^p = 1 - \exp \left(-\beta \left(\frac{(d_i^p - y_i^p)}{(y_i^p(\text{máximo}) - y_i^p(\text{mínimo}))} \right)^2 \right) \quad (2.3.49)$$

$$\delta_i^p = \text{sgn} (d_i^p - y_i^p) * E^p \quad (2.3.50)$$

$$\delta_k^p = \sum w_{ik}^p * \delta_i^p \quad (2.3.51)$$

onde:

E^p = erro no padrão p ;

sgn = função sinal;

$d_i^p - y_i^p$ = diferença entre a saída desejada (d_i^p) e a saída real (y_i^p) do neurônio i para cada padrão p ;

β = é um fator escalar;

δ_i^p = erro da unidade de saída i para o padrão p ;

δ_k^p = erro da unidade intermediária k para o padrão p .

Alguns parâmetros utilizados pela arquitetura NEFCLASS, incluem:

- unidades de entrada: correspondem às variáveis e devem ser iguais ao número de características dos padrões (componentes de um vetor) a ser classificado;
- regras: são criadas durante o algoritmo de aprendizado ou podem ser introduzidas e/ou alteradas. Existe um número máximo de (500) regras; caso o número seja muito extenso, pode comprometer a velocidade de processamento;
- saídas: o número máximo de classes de saída é de 50 e devem ser iguais ao número de classes padrões utilizadas na fase de treinamento.

Os nomes das classes devem ser definidos individualmente para as unidades de saída. Para cada unidade de entrada um número de CDs são definidos, sendo utilizadas funções de pertinência triangulares. Estes conjuntos possuem algumas características:

- cada variável pode assumir o mesmo número de CDs;
- cada unidade de entrada pode ter no máximo (7) CDs.

O NEFCLASS está disponível através da URL: <http://fuzzy.cs.uni-magdeburg.de/nefclass/>.

As pesquisas mais recentes em SNDs concentram-se na determinação automática do número de termos lingüísticos na etapa de fuzzificação das variáveis, resultando diretamente no número de regras neuro-fuzzy (Rezende et al., 2003).

2.3.6 Considerações Finais

Esta seção apresentou algumas arquiteturas de SNDs aplicadas no treinamento híbrido envolvendo RNAs e a LD. Nas arquiteturas descritas, as RNAs e a LD trabalham juntas de forma a extrair o que há de melhor em cada uma. As arquiteturas mostram diferentes formas de aprendizagem utilizadas na implementação dos SNDs com o objetivo de aumentar o desempenho dos algoritmos de aprendizagem.

Tendo como base o referencial acima, propõe-se nesta tese um sistema híbrido neuro-difuso explorando como metodologia de aprendizagem a implementação de diferentes operações aritméticas difusas *t-normas* e *t-conormas*. O SND proposto utiliza uma interface difusa, a qual fornece um vetor de entrada para uma RNA multicamadas com o operador lógico *E* na camada intermediária e o operador lógico *OU* na camada de saída da RNA.

No próximo capítulo são descritos os modelos de aprendizagem investigados.

Capítulo 3

Modelos de Aprendizagem

Este capítulo descreve os modelos de aprendizagem a partir das arquiteturas investigadas. Inicialmente, faz-se uma descrição das arquiteturas de SNDs com suas respectivas operações aritméticas difusas e, após, são apresentadas as metodologias de aprendizagem incluindo os diferentes tipos de neurônios difusos.

3.1 Considerações Iniciais

O objetivo desta tese é avaliar os modelos de aprendizagem aplicados aos SNDs explorando diferentes operações aritméticas difusas através de uma metodologia de aprendizagem. No desenvolvimento da metodologia de aprendizagem são analisados os desempenhos das arquiteturas NEFCLASS por meio das operações aritméticas difusas padrões *max/min* e do SisND/EPs, utilizando os neurônios difusos propostos por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994), exemplificados no Quadro 3.1.1.

Quadro 3.1.1 - Arquiteturas de SNDs e as operações aritméticas difusas

Arquiteturas de SNDs			Operações Aritméticas Difusas
NEFCLASS			<i>Max/Min</i>
SisND/EPs	N e u r ô n i o s	Hirota e Pedrycz	Soma/Produto Algébrico OU (E)
			Soma/Produto de Hamacher OU (E)
			Soma/Produto de Einstein OU (E)
			<i>Max/Min</i>
	D i f u s o s	Kwan e Cai	Soma/Produto Algébrico OU (E)
			Soma/Produto de Hamacher OU (E)
			Soma/Produto de Einstein OU (E)
			<i>Max/Min</i>

De acordo com o Quadro 3.1.1, a primeira arquitetura (NEFCLASS) investiga as operações aritméticas difusas padrões max/min ; a segunda arquitetura, o SisND/EPs explora, além dos neurônios difusos propostos por Hirota-Pedrycz e Kwan-Cai, as operações aritméticas difusas t -normas e t -conormas: Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein OU(E), juntamente com as operações aritméticas difusas padrões max/min .

Nesta tese, a arquitetura NEFCLASS foi o modelo padrão implementado porque representa as operações aritméticas difusas padrões max/min , as quais formam a base de comparação com as demais operações aritméticas difusas t -normas e t -conormas.

A seguir são descritas as etapas utilizadas na implementação:

Etapas_1 – inicialmente, foram investigadas as operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E) e Soma/Produto de Einstein OU(E), utilizando o neurônio difuso proposto por Hirota-Pedrycz juntamente com as operações aritméticas difusas padrões max/min por meio do SisND/EPs;

Etapas_2 – a seguir, foram exploradas as operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E) e Soma/Produto de Einstein OU(E), utilizando o neurônio difuso proposto por Kwan-Cai e com as operações aritméticas difusas padrões max/min através do SisND/EPs;

Etapas_3 – os resultados obtidos nas etapas 1 e 2 foram comparados com os resultados obtidos com as operações aritméticas difusas padrões max/min , implementados por meio da arquitetura NEFCLASS.

Na seqüência descrevem-se as etapas utilizadas na implementação das metodologias de aprendizagem.

3.2 Metodologias de Aprendizagem

O processo de aprendizagem das arquiteturas consiste em atualizar os pesos sinápticos utilizando as operações aritméticas difusas Produto Algébrico, Produto de Hamacher, Produto de Einstein, juntamente com as operações aritméticas difusas padrões min para os neurônios E . Nos neurônios OU são utilizadas as operações aritméticas difusas Soma Algébrica, Soma de Hamacher, Soma de Einstein e as operações aritméticas difusas padrões max na etapa de avaliação do conjunto de

regras. Esta etapa está dividida nas fases *feedforward* e *backward*, baseadas no algoritmo de retropropagação modificado¹⁸.

Na seqüência faz-se a descrição resumida dos modelos de aprendizagem para as operações aritméticas difusas *t-normas* e *t-conormas*.

3.2.1 Aprendizagem com o Neurônio Difuso Hirota-Pedrycz (1992, 1994)

O principal elemento dos modelos de aprendizagem consiste na implementação do neurônio difuso. A característica do neurônio difuso proposto por Hirota-Pedrycz (1992, 1994) consiste dos neurônios difusos *E* e *OU*, descritos a seguir.

Neurônio Difuso E:

A saída do sinal é implementada aplicando a função de ativação logística.

Os valores de x_j e v_{kj} são combinados por uma operação *t-conorma* S para produzir p_k^p :

$$p_k^p = S(v_{kj}^p, x_j^p), j = 1, \dots, n \quad (3.2.1)$$

onde:

p_k^p = resultado da operação aritmética difusa *t-conorma* S aplicada no neurônio k da camada intermediária para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p , onde $v_{kj}^p \in [-0.5, 0.5]$;

A entrada da informação p_k^p é agregada por uma operação aritmética difusa *t-norma* T para produzir a saída net_k^p :

$$net_k^p = \theta_k^p + T(S(p_{11}^p), \dots, S(p_{kj}^p), \dots, S(p_{qn}^p)) \quad (3.2.2)$$

$$z_k^p = \varphi(net_k^p) \quad (3.2.3)$$

onde:

net_k^p = entrada da RNA atingida pelo neurônio k da camada intermediária para o padrão p ;

$\varphi(net_k^p)$ = função de ativação aplicada na entrada da RNA atingida pelo neurônio k da camada intermediária da RNA para o padrão p ;

θ_k^p = *bias* da camada intermediária k para o padrão p ;

z_k^p = resultado de saída da camada intermediária k para o padrão p ;

T = operação aritmética difusa *t-norma* (Produto Algébrico, Produto de Hamacher, Produto de

¹⁸ A nomenclatura utilizada no algoritmo de retropropagação é baseada no trabalho de Fausett (1994, p. 289).

Einstein ou a operação aritmética difusa padrão *min*);

S = operação aritmética difusa *t-conorma* (Soma Algébrica, Soma de Hamacher, Soma de Einstein ou a operação aritmética difusa padrão *max*).

Neurônio Difuso OU:

A saída do sinal é implementada aplicando a função de ativação logística.

Os valores z_k^p e w_{ik}^p são combinados por uma operação *t-norma* T para produzir p_i^p :

$$p_i^p = T(w_{ik}^p, z_k^p), k = 1, \dots, q \quad (3.2.4)$$

onde:

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p .

A entrada da informação p_i^p é agregada por uma operação *t-conorma* S para produzir a saída net_i^p :

$$net_i^p = \theta_i^p + S(T(p_{11}^p), \dots, T(p_{ik}^p), \dots, T(p_{mq}^p)) \quad (3.2.5)$$

$$y_i^p = \varphi(net_i^p) \quad (3.2.6)$$

onde:

θ_i^p = *bias* da camada de saída i para o padrão p ;

$(p_{11}^p), \dots, (p_{ik}^p), \dots, (p_{mq}^p)$ = resultado da operação *t-norma* T aplicada no neurônio i da camada de saída para o padrão p ;

net_i^p = entrada da RNA atingida pelo neurônio i da camada de saída para o padrão p ;

$\varphi(net_i^p)$ = função de ativação logística aplicada na entrada da RNA atingida pelo neurônio i da camada de saída da RNA para o padrão p .

3.2.2 Aprendizagem com o Neurônio Difuso Kwan-Cai (1994)

O neurônio difuso proposto por Kwan-Cai (1994) consiste dos neurônios difusos E e OU , descritos a seguir.

Neurônio Difuso E:

A saída do sinal é implementada aplicando a função de ativação logística.

O neurônio difuso E da camada intermediária k apresenta o sinal de entrada x_j interagindo com o peso sináptico v_{kj} para produzir o produto p_k^p :

$$p_k^p = (v_{kj}x_j), j = 1, \dots, n \quad (3.2.7)$$

A entrada da informação p_k^p é agregada por uma operação aritmética difusa *t-norma* T para computar a saída net_k^p :

$$net_k^p = \theta_k^p + T\{(p_{11}^p), \dots, (p_{kj}^p), \dots, (p_{qn}^p)\} = \theta_k^p + T\{v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n\} \quad (3.2.8)$$

onde:

θ_k^p = *bias* da camada intermediária k para o padrão p ;

p_k^p = resultado da operação aritmética *produto* aplicada no neurônio k da camada intermediária para o padrão p ;

$\varphi(net_k^p)$ = função de ativação aplicada na entrada da RNA atingida pelo neurônio k da camada intermediária da RNA para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p ;

x_j^p = valor de entrada *fuzzificado* j para o padrão p .

A saída do neurônio z_k^p é computada por:

$$z_k^p = \varphi(net_k^p - \theta_k^p) = \varphi(T\{v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n\} - \theta_k^p) \quad (3.2.9)$$

onde:

z_k^p = resultado da saída da camada intermediária k para o padrão p ;

Neurônio Difuso OU:

A saída do sinal é implementada aplicando a função de ativação logística.

No neurônio difuso *OU* o sinal de entrada z_k^p interage com o peso sináptico w_{ik} para produzir o produto p_k^p :

$$p_k^p = w_{ik}z_k, k = 1, \dots, q \quad (3.2.10)$$

A entrada da informação p_k^p é agregada por uma operação *t-conorma* S para produzir a saída net_i^p :

$$net_i^p = \theta_i^p + S\{(p_{11}^p), \dots, (p_{ik}^p), \dots, (p_{mq}^p)\} = \theta_i^p + S\{w_{11}z_1, \dots, w_{ik}z_k, \dots, w_{mq}z_q\} \quad (3.2.11)$$

onde:

θ_i^p = *bias* da camada de saída i para o padrão p ;

p_i^p = resultado da operação aritmética *produto* aplicada no neurônio i da camada de saída para o padrão p ;

$\varphi(net_i^p)$ = função de ativação logística aplicada na entrada da RNA atingida pelo neurônio i da camada de saída da RNA para o padrão p ;

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da

camada de saída no padrão p ;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

A saída do neurônio y_i^p é computada por:

$$y_i^p = \varphi(\text{net}_i^p - \theta_i^p) = \varphi(S\{w_{11}z_1, \dots, w_{ik}z_k, \dots, w_{mq}z_q\} - \theta_i^p) \quad (3.2.12)$$

onde:

y_i^p = camada de saída i para o padrão p .

3.2.3 Aprendizagem com as Operações Aritméticas *Max* e *Min* (Nauck, 1994)

Na implementação das operações aritméticas difusas padrões *max/min* (por não serem deriváveis) o SisND/EPs não utilizou o método do gradiente descendente, e, sim, a função sinal do gradiente (*sgn*), como sugerido por Nauck (1994). A seguir a descrição do neurônio difuso *E* e *OU*:

Neurônio Difuso E:

Os valores x_j^p e v_{kj}^p são combinados por uma operação *t-norma* (*min*) para produzir p_k^p :

$$p_k^p = \min(v_{kj}^p, x_j^p) \quad (3.2.13)$$

onde:

p_k^p = resultado da operação *t-norma* (*min*) aplicada no neurônio k da camada intermediária para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p , onde $v_{kj}^p \in [-0.5, 0.5]$;

À entrada da informação p_k^p são acrescentados o *bias* e seu respectivo peso para produzir a saída net_k^p :

$$\text{net}_k^p = \theta_k^p + ((p_{k1}^p), \dots, (p_{kj}^p), \dots, (p_{kn}^p)) \quad (3.2.14)$$

$$z_k^p = \varphi(\text{net}_k^p) \quad (3.2.15)$$

onde:

net_k^p = entrada da RNA atingida pelo neurônio k da camada intermediária para o padrão p ;

$\varphi(\text{net}_k^p)$ = função de ativação aplicada na entrada da RNA atingida pelo neurônio k da camada intermediária da RNA para o padrão p ;

θ_k^p = *bias* da camada intermediária k para o padrão p ;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

Neurônio Difuso OU:

Os valores z_k^p e w_{ik}^p são combinados por uma operação *t-conorma* (*max*) para produzir p_i^p :

$$p_i^p = \max (w_{ik}^p, z_k^p) \quad (3.2.16)$$

onde:

p_i^p = resultado da operação *t-conorma* (*max*) aplicada no neurônio *i* da camada de saída para o padrão *p*;

w_{ik}^p = intensidade da conexão do neurônio *k* da camada intermediária para o neurônio *i* da camada de saída no padrão *p*.

À entrada da informação p_i^p são acrescentados o *bias* e seu respectivo peso para produzir a saída net_i^p :

$$net_i^p = \theta_i^p + ((p_{1k}^p), \dots, (p_{ik}^p), \dots, (p_{mk}^p)) \quad (3.2.17)$$

$$y_i^p = \varphi(net_i^p) \quad (3.2.18)$$

onde:

θ_i^p = *bias* da camada de saída *i* para o padrão *p*;

$(p_{1k}^p), \dots, (p_{ik}^p), \dots, (p_{mk}^p)$ = resultado da operação aritmética difusa *t-norma T* aplicada no neurônio *i* da camada de saída para o padrão *p*;

net_i^p = entrada da RNA atingida pelo neurônio *i* da camada de saída da RNA para o padrão *p*;

$\varphi(net_i^p)$ = função de ativação logística aplicada na entrada da RNA atingida pelo neurônio *i* da camada de saída da RNA para o padrão *p*.

O critério de convergência do algoritmo de retropropagação foi baseado no erro quadrático total $E(t)$, ou seja, se o $E(t) \leq 0.40$, o registro apresenta a tendência para um ponto comum, ou seja, convergiu, caso contrário, o registro não apresenta a tendência para um ponto comum. Este critério foi adotado tendo em vista que a arquitetura NEFCLASS utiliza o mesmo critério de convergência, e, ambos os algoritmos são comparados.

As metodologias de aprendizagem incluindo os algoritmos completos de treinamento encontram-se no Apêndice 4.

Na seqüência faz-se a descrição de como será contornado o problema da não-derivabilidade das operações aritméticas difusas padrões *max/min*, quando utilizado o algoritmo de retropropagação modificado.

3.2.4 Restrições Referente às Operações Aritméticas *Max/Min*

As operações aritméticas difusas padrões *max/min*, podem apresentar pontos de descontinuidade, sendo funções não-deriváveis. Essas funções possuem valores impróprios em determinados pontos, conduzindo os valores *fuzziificados* de entrada x_j ao ∞ e para h tendendo a zero, ou seja, não existe derivada no ponto x_j , de acordo com a Equação:

$$\varphi'(x_j) = \lim_{h \rightarrow 0} \frac{f(x_j + h) - f(x_j)}{h} = \infty \quad (3.2.19)$$

A partir da Equação 3.2.19, se aplicado o algoritmo de treinamento de retropropagação com as operações aritméticas difusas padrões *max/min*, pode ocorrer o seguinte (Brasil, 1999):

Se, no momento do treinamento, o algoritmo não encontrar pontos de descontinuidade, dependendo do intervalo que esteja sendo aplicado, pode-se dizer que a RNA tem possibilidades de convergir; caso se verifiquem pontos de descontinuidade, ocorrerá que neste ponto não existirá derivada, fazendo com que a RNA tenha possibilidades de não convergir. (p. 158).

Nesta tese, na implementação das operações aritméticas difusas padrões *max/min* não foi utilizado o método do gradiente descendente. O método escolhido para a implementação utilizou apenas a função sinal do gradiente (Nauck, 1994)¹⁹. Os resultados desta implementação, incluindo as operações aritméticas difusas padrões *max/min*, foram comparados com os resultados obtidos com as demais operações aritméticas difusas *t-normas* e *t-conormas*, de acordo com o Quadro 3.1.1 descrito no capítulo 3, Modelos de Aprendizagem.

Na implementação das operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E) e Soma/Produto de Einstein OU(E) foi aplicada à função logística numa etapa anterior à etapa de derivação.

3.3 Originalidade da Pesquisa

A originalidade da pesquisa concentra-se na exploração de diferentes modelos de aprendizagem aplicados aos SNDs com o objetivo de encontrar um SND com melhores percentuais de acertos. Estes modelos de aprendizagem foram avaliados no diagnóstico dos *EEs* e dos *ENEs*.

¹⁹

A função sinal do gradiente encontra-se descrita no Apêndice 4, Algoritmos de Aprendizagem.

Para explorar os modelos de aprendizagem foram utilizados os paradigmas:

- (1) avaliação e validação de diferentes tipos de neurônios difusos;
- (2) análise dos operadores lógicos OU(E) na camada intermediária e de saída do SND;
- (3) avaliação e validação das operações aritméticas difusas *t-normas* e *t-conormas*:
 - *Max/min*;
 - Soma/Produto Algébrico OU(E);
 - Soma/Produto de Hamacher OU(E);
 - Soma/Produto de Einstein OU(E).

No tratamento das informações imprecisas, aplicada na área de domínio dos *EEs* e os *ENEs*, foi utilizada a LD por meio de diferentes operações aritméticas difusas.

Até o momento não foram encontradas referências na literatura consultada relacionadas com a exploração das diferentes operações aritméticas difusas propostas nesta tese, bem como técnicas para tratar as informações *imprecisas* relacionadas com a área de domínio explorada na tese. Na literatura foram encontradas, além da arquitetura NEFCLASS, pesquisas envolvendo neurônios de referência (Pedrycz e Rocha, 1993; Pedrycz, 1993, 2004) e pesquisas com as operações aritméticas difusas Lukasiewicz; Produto Limitado e *t-norma* de Yager, além do produto algébrico e da operação *min* (Pedrycz et al., 2006).

Portanto, nesta tese, faz-se uma comparação original dos modelos de aprendizagem utilizando diferentes operações aritméticas difusas e diferentes tipos de neurônios difusos, aplicados num domínio da área médica.

3.4 Considerações Finais

Este capítulo apresentou os modelos de aprendizagem das arquiteturas NEFCLASS, descrevendo as operações aritméticas difusas padrões *max/min* e o SisND/EPs, o qual utiliza os neurônios difusos propostos por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994), aplicados na resolução de problemas de diagnóstico médico, especificamente no domínio da neurologia.

No próximo capítulo faz-se a descrição do domínio de aplicação, área escolhida para avaliar e validar os modelos de aprendizagem.

Capítulo 4

Domínio de Aplicação

Este capítulo aborda o domínio de aplicação dos modelos de aprendizagem relacionados com a área da neurologia, especificamente no Diagnóstico Diferencial de Eventos Paroxísticos (EPs), envolvendo os *EEs* e os *ENEs*. A escolha por esta área de pesquisa está relacionada com: (1) dificuldade em diferenciar um *EE* de um *ENE* devido à inexistência de limites precisos entre essas categorias diagnósticas; (2) a área da epilepsia já vinha sendo estudada pelo grupo de Engenharia Biomédica da Universidade Federal de Santa Catarina/SC e, por esta razão, a disponibilidade de informações tornou possível a realização desta tese. Buscam-se, pela exploração das diferentes operações aritméticas difusas, resultados mais precisos.

4.1 Considerações Iniciais

John Hughins Jackson, no período de 1835 a 1911, definiu a epilepsia como uma descarga neuronal paroxística²⁰ que leva à convulsão generalizada, quando o tecido normal é invadido pela atividade da crise (Silva, 1987).

Em razão da falta de uniformidade na classificação das epilepsias, a *International League Against Epilepsy* (ILAE) instaurou um comitê internacional para tentar padronizá-la, procurando critérios nas informações que podem ser obtidas dos pacientes (anamnese) e nas características do eletroencefalograma (EEG)²¹ (Guerreiro et al., 2000).

A maior contribuição da ILAE foi o estabelecimento de terminologias e classificações padronizadas para diagnosticar crises e síndromes epiléticas, fornecendo um vocabulário universal para facilitar a comunicação entre os médicos e, também, estabelecer uma taxonomia para execução das pesquisas em epilepsia (Ilae, 2006).

As duas últimas décadas foram marcadas por um interesse crescente nos chamados *ENEs*. Os neuropsicólogos e psiquiatras descobriram nos *ENEs* um campo fértil para o estudo dos distúrbios

²⁰ O termo *paroxístico* significa início e término abrupto de um evento epilético.

²¹ Registro gráfico da atividade elétrica encefálica, obtido mediante o uso de eletroencefalógrafo (Ferreira, 2004).

de personalidade, associados ou não às epilepsias (Costa et al., 1988). As conseqüências do não-reconhecimento dos *ENEs* podem ser catastróficas, pois pacientes não epiléticos podem ser submetidos, erroneamente, a regimes terapêuticos com anticonvulsivantes. O verdadeiro desafio diagnóstico e de conduta são os chamados "pacientes mistos", portadores de *EEs* e de *ENEs*. O volume deste último contingente não pode ser desprezado uma vez que representa de 10% a 60% da população atendida em centros de epilepsia tradicionais (Gates, 1993_a).

Crises não-epilépticas, também chamadas de "crises pseudoepilépticas", "histéricas" ou "psicogênicas", são mudanças paroxísticas no comportamento e assemelham-se a crises epilépticas, porém não estão associadas a alterações quantificáveis na atividade elétrica do cérebro (Conway et al., 2006).

Outros termos encontrados na literatura citam a "pseudocrise", "epilepsia histórica" ou "histeroepilepsia", porém, de acordo com o autor, tais termos não são adequados por denotarem uma noção de falsidade ou engano a um episódio que é, de fato, verdadeiro, porém não epiléptico (Costa et al., 1998).

Nesta tese são utilizados os termos *EEs* e *ENEs*, representando, dessa forma, as várias terminologias encontradas na literatura.

Na seqüência, descrevem-se a classificação e as características relacionadas aos *EEs* e *ENEs*.

4.2 Classificação dos Eventos Epilépticos

O primeiro sistema para classificação das convulsões apresentado pela ILAE foi em 1969, composto por quatro divisões: crises parciais, crises generalizadas, crises unilaterais e crises não classificadas. O segundo sistema é a classificação clínica e eletroencefalográfica de crises epilépticas (Ilae, 1981). A terceira classificação sugerida pela ILAE foi publicada em 1985 e revisada em 1989, a qual auxilia na classificação de síndromes epilépticas e epilepsias (Ilae, 1989). A proposta de 2001, ainda em estudos sugerida pela ILAE, visa a uma nova metodologia de classificação com o objetivo de ser mais sistêmica, detalhada e precisa (Ilae, 2001)²². Tanto a classificação de 1981 como a de 1989 são as mais utilizadas e as que atualmente persistem na área médica.

Nesta tese foi adotada a classificação de 1981, pois a pesquisa envolve diagnóstico diferencial de crises epilépticas.

²² Este estudo está sendo formalizado e debatido entre vários pesquisadores da área. Alguns artigos com as idéias de cada autor são encontrados em (Avanzini, 2003; Berger e Blackstone, 2003; Engel, 2003; Fisher, 2003; Lüders, Najm e Wyllie, 2003; Wolf, 2003).

Na seqüência, faz-se uma descrição resumida das crises parciais, crises generalizadas e crises não-classificáveis sugerida pela ILAE.

4.2.1 International League Against Epilepsy (Ilae, 1981)

A classificação das crises epiléticas sugerida pela ILAE está dividida em:

1. Crises parciais (ou focais):

Crise Parcial Simples (CPS) - sem prejuízo da consciência:

- com sinais motores
- com sinais sensitivos, somatossensoriais ou especiais
- com sinais ou sintomas autonômicos
- com sinais psíquicos

Crise Parcial Complexa (CPC) - com prejuízo da consciência

- início de CPS seguida por alteração da consciência
- alteração da consciência no início

Crises secundariamente generalizadas:

- CPS evoluindo para Crises Generalizadas Tônico-Clônicas (CGTC)
- CPC evoluindo para CGTC
- CPS evoluindo para CPC e, então, para CGTC

2. Crises generalizadas (desde o início) - com prejuízo da consciência:

CGTC

Crise de ausência

Crise de ausência atípica

Crises mioclônicas

Crises tônicas

Crises clônicas

Crises atônicas

3. Crises não-classificáveis (informações incompletas e/ou inadequadas).

4.2.2 Características dos Eventos Epiléticos

A incidência da epilepsia varia de acordo com a localização geográfica, mas ocorre com maior freqüência nos países em desenvolvimento, onde há mais desnutrição, doenças infecciosas

e deficiência no atendimento médico. Em países mais desenvolvidos, a incidência é de aproximadamente 1%, subindo para 2% em nações menos desenvolvidas (Epilepsia, 2006).

Epilepsia e as doenças mentais são conhecidas desde a Antigüidade. A história das duas doenças é repleta de associações com a religião e a superstição. Com o passar do tempo, os médicos consideraram os fenômenos epiléticos como psiquiátricos, e pacientes com epilepsia eram freqüentemente internados em manicômios (Engel, 2006).

Jean Martin Charcot e Willian Gowers, na segunda metade do século XIX, foram os primeiros a discernir, com algum grau de eficiência, características de EPs, lembrando epilepsia e histeria (Ramani, 1993).

Admitem-se os EEs como um grupo de doenças que têm em comum crises epiléticas, as quais são eventos clínicos que refletem uma disfunção temporária de um conjunto de neurônios de parte do encéfalo caracterizando as crises focais ou de uma área mais extensa, envolvendo simultaneamente os dois hemisférios cerebrais, dando origem às crises generalizadas (Guerreiro et al., 2000).

Se as alterações epiléticas permanecerem restritas a uma parte do cérebro, a crise chama-se "parcial"; se o cérebro inteiro está envolvido, chama-se "generalizada". Crises parciais simples não ocasionam a perda da consciência e caracterizam-se por distorções na percepção auditiva ou visual, desconforto estomacal, sensação súbita de medo e/ou movimentos estranhos de uma parte do corpo. Crises parciais complexas são crises que, como as parciais simples, iniciam-se num foco determinado no cérebro, mas espalham-se para outras áreas, causando perturbações da consciência. A pessoa fica confusa e pode caminhar sem rumo, falar sem coerência, salivar em excesso, morder a língua e realizar automatismos, como puxar a roupa ou virar a cabeça de um lado para outro repetidas vezes (Epilepsia, 2006).

Entre as causas que dificultam o tratamento das epilepsias existe a possibilidade do falso diagnóstico, ou seja, o resultado da incorreta identificação dos tipos de crises, como, por exemplo, as ausências simples confundidas com *EEs* parciais complexos (pseudo-ausências), tratadas com medicação inadequada.

De acordo com a descrição sugerida pela ILAE essas crises apresentam os seguintes sintomas (ILAE, 1981; Tierney, McPhee e Papadakis, 2004):

Crise Parcial Simples

· Sintomas motores: atividades motoras involuntárias, movimentos da cabeça para um lado, perda da habilidade de falar, sons vocais;

- Sintomas autonômicos: vômito, palidez, rubor/vermelhão, dilatação das pupilas, incontinência urinária, transpiração excessiva, sialorréia (baba);
- Sintomas somatossensoriais: sensação de formigamento, sentimento de dormência, surgimento de flashes, pessoas e figuras; zunido, odor e sabor desagradável, sensação de queda no espaço;
- Sintomas psíquicos: perda da habilidade de falar, memória repentina de um fato passado, falar que conhece alguém ou lugar sem nunca ter conhecido ou estado lá; falar que jamais compreendeu alguma coisa ou que compreende algo; estados de sonhos, desinteresse, terror, sensação de extremo prazer ou desprazer, medo, intensa depressão, raiva ou fúria, pupilas dilatadas, palidez, rubor/vermelhão, palpitação, ataques de riso (gargalhada), distorção no tamanho dos objetos (macropsia/micropsia), distorção de sons, surgimento de ruídos rápidos, despersonalização das pessoas (alteração no tamanho de partes do corpo).

Crise parcial complexa

Atividades motoras involuntárias seguidas de eventos de anamnésia, apresentando os automatismos: alimentares, arremedação/mímica, gesticulares e verbais. O paciente apresenta a tendência de arrumar várias vezes suas próprias roupas.

As crises parciais são aquelas nas quais, em geral, as primeiras manifestações clínicas e eletroencefalográficas indicam ativação de um sistema neuronal limitado à parte de um hemisfério cerebral (Guerreiro et al., 2000). O que distingue a CPS da CPC é o comprometimento da consciência nesta última. A consciência, aqui, é entendida como a capacidade de percepção consciente e de interação com o ambiente de forma adequada. Diz-se que há comprometimento da consciência quando essa capacidade estiver prejudicada. É comum durante uma CPC o indivíduo ser capaz de perceber e interagir, porém de forma inadequada.

Crises generalizadas

- Crise tônico-clônica: súbita contração tônica dos músculos, aparecimento de som agudo/áspero, gemido, grito, deixando o paciente cair em estado clônico. Contorções do corpo para trás, inibição da respiração, cianose, sabor amargo na língua, liberação esfinteriana, grunidos respiratórios, sialorréia, respiração profunda e relaxamento muscular. Após a crise, o paciente sente dores no corpo, dor de cabeça e pode cair em sono profundo;
- Crise de ausência: surgimento de automatismos de atenção, mímica com gestos de arrumar a roupa e olhar fixo no espaço vazio. Breves interrupções, entre elas rotação dos olhos para cima, voz, caminhada, alimentares e interrupções na escuta;

- Crise de ausência atípica: com suaves movimentos clônicos da pálpebra e canto da boca. Os objetos são derrubados, o paciente pode apresentar: (1) sinais atônicos: diminuição do tonus muscular, tristeza, queda do corpo, caminhar de braços abertos; (2) sinais tônicos: contração da musculatura, contorcimento da cabeça para trás ou para os lados, corpo arcado e (3) automatismos: caminhar vagarosamente, ranger os lábios e resmungo;
- Crise mioclônica: súbitas e breves contrações de choque generalizadas pelo corpo ou em uma parte da face, podendo ocorrer durante o sono ou em vigília, repetidamente ou isoladas;
- Crise tônica: rígida contração muscular, virar os olhos, cabeça torcida para trás ou para os lados, rotação do corpo, palidez facial e cianose;
- Crise clônica: caracterizada por crises clônicas repetitivas, ou seja, o paciente deixa-se cair em estado clônico;
- Crise atônica: caracterizada por crises breves, diminuição do tônus muscular, queda do corpo ou cabeça para um lado e queda brusca.

4.3 Sistematização dos Eventos Não Epilépticos

A consequência natural da sistematização no diagnóstico de *ENEs* encontra-se na sua classificação em origem fisiológica e psicogênica (Costa et al., 1998), como mostra a Figura 4.1.1.

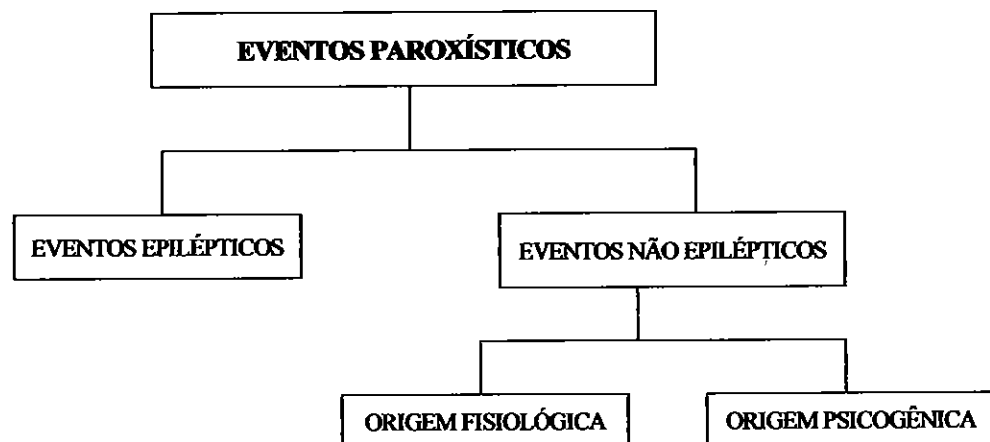


Figura 4.1.1 - Classificação dos eventos paroxísticos (adaptado de Costa et al. (1998))

4.3.1 Classificação dos Eventos Não Epilépticos

A classificação dos *ENEs* é definida como:

1. *ENE* de origem fisiológica:

- Síncope
- Ataque isquêmico transitório
- Distúrbios de movimento paroxísticos
- Distúrbios tóxicos paroxísticos
- Alucinações orgânicas não tóxicas
- Distúrbios do sono
- Paroxismos decorrentes de insultos neurológicos agudos
- Distúrbios endócrinos paroxísticos
- Mioclonia não epiléptica

2. *ENE* de origem psicogênica:

Transtornos com sintomas psicóticos

Padrão comportamental reforçado

Transtornos de ansiedade:

- Transtorno do pânico com agorafobia
- Transtorno do pânico sem agorafobia
- Transtorno do estresse pós-traumático
- Transtorno do estresse agudo

Transtornos dissociativos:

- Fuga dissociativa
- Transtorno de despersonalização
- Transtorno dissociativo não especificado

Transtornos somatoformes:

- Transtorno de somatização
- Transtorno somatoforme indiferenciado
- Transtorno conversivo (crise histérica)
- Hipocondria
- Transtorno somatoforme não especificado

Simulação

Transtornos factícios.

4.3.2 Características dos Eventos Não Epilépticos

Numa visão mais psicológica, os *ENEs* são classificados como um processo de somatização no qual o indivíduo faz uso do corpo, através de queixas e/ou sintomas físicos, como forma de expressar as suas angústias (Santo, Maineri e Portuquez, 2004).

Luther et al. (1982) e, posteriormente, Gates (1986) classificaram de forma funcional os *ENEs* de natureza psicogênica. A partir dessas observações, foi proposta uma subdivisão dos *ENEs* psicogênicos em cinco grupos, descritos a seguir:

1) grupo de pacientes com dificuldade de adaptação: tipicamente pacientes que utilizam seus *ENEs*, objetivando o controle do ambiente;

2) grupo com erro de interpretação dos seus próprios sintomas: este grupo é caracterizado por pacientes que interpretam qualquer sensação corporal como um *EE* e relatam o episódio como tal;

3) grupo de pacientes psicóticos: pacientes com quadro psicótico, mas com comportamentos ocasionalmente lembrando estados confusionais próprios de *EEs*;

4) grupo com conflitos emocionais: maior de todos os subgrupos, inclui todas as reações do tipo conversivas; tipicamente, os episódios são desencadeados a partir de conflitos emocionais e expressam-se por *EEs*; são comuns as histórias de desajuste da estrutura familiar, abuso sexual e traumas emocionais múltiplos;

5) grupo de pacientes com tendência à supervalorização dos seus sintomas: caracteriza-se por pacientes portadores de CPS, as quais são seguidas por um *ENE*, desencadeados pela valorização dos sintomas.

Os tipos mais frequentes de *ENEs* correspondem às crises histéricas denominadas "crise convulsiva conversiva" (Apa, 1994; Who, 1993). Por esse motivo, faz-se uma descrição dos seus principais sintomas, os quais são utilizados na implementação do SisND/EPs.

A histeria baseia-se em modificações fisiológicas do sistema nervoso central. Sua essência deve ser expressa numa fórmula que leve em consideração as condições de excitabilidade nas diferentes partes do sistema nervoso central. Charcot sustenta com firmeza a opinião de que a histeria é um quadro bem definido, que pode ser reconhecido com bastante clareza nos casos extremos daquilo que se conhece como "grande histeria" ou "histeroepilepsia" (Freud, 1998).

A sintomatologia da grande histeria é extremamente rica e compõe-se de uma série de sintomas (Freud, 1998):

- Ataques convulsivos: são precedidos de uma aura: pressão no epigástrico, compressão na garganta, latejamento das têmporas, zunido nos ouvidos, espasmos da faringe, como se uma bola estivesse subindo do epigástrico para a garganta;

- Zonas histerogêneas: referem-se às áreas supersensíveis do corpo, nas quais um leve estímulo desencadeia um ataque cuja aura, muitas vezes, inicia por uma sensação proveniente desta área;

- Distúrbios de sensibilidade: estes são os sinais mais freqüentes das neuroses e os mais importantes do ponto de vista diagnóstico. Constituem-se em anestesia ou hiperanestesia, as quais podem afetar a pele, as membranas mucosas, os ossos, os músculos e nervos, os órgãos dos sentidos e os intestinos;

- Distúrbios da atividade sensorial: estes distúrbios podem afetar todos os órgãos dos sentidos e seus sintomas característicos apresentam: fundo de olho normal ao exame, ausência do reflexo conjuntival (reflexo corneano diminuído), estreitamento do campo visual, redução da percepção luminosa e acromaptosia, isto é, o paciente enxerga na cor preta, cinza ou branca;

- Paralisias: as paralisias histéricas são mais raras do que a anestesia e quase sempre acompanhadas de anestesia da parte do corpo paralisada. Não existe paralisia facial histérica; quando muito, além da paralisia dos membros, pode haver uma contratura dos músculos faciais e da língua. A paralisia pode afetar segmentos do corpo - mão, ombro, cotovelo, etc.;

- Contraturas: ato ou efeito de contrair-se. Apresentam uma persistência incomum, não relaxam com o sono, como ocorre com as contraturas orgânicas. Somente recobram sua intensidade total quando o paciente acorda.

Duarte (1988) descreve uma categoria que agrupa, de um lado, uma sintomatologia física (dores de cabeça, insônia, desmaio, falta de apetite, tremores, etc.) e, de outro, sintomas morais (irritação, aborrecimento, depressão, indisposição tristeza, fúria, etc.).

4.4 Considerações Finais

Este capítulo descreveu, resumidamente, alguns tópicos relacionados aos *EEs* e *ENEs*. As variáveis de entrada implementadas no SisND/EPs referentes aos *EEs* integram as crises parciais simples, crises parciais complexas e crises generalizadas. Na classificação dos *ENEs* são trabalhados os sintomas de origem fisiológica e de origem psicogênica. Os tipos mais freqüentes de *ENEs* correspondem às crises histéricas; por esse motivo, faz-se uma descrição dos seus principais sintomas, os quais são utilizados na implementação do SisND/EPs.

Nesta tese, para caracterizar um *EE* de um *ENE*, são utilizadas quatorze variáveis de entrada, as quais são descritas no Capítulo 5, modelos de aprendizagem. As variáveis de entrada foram definidas juntamente com o médico neurologista que acompanhou o desenvolvimento da pesquisa.

No próximo capítulo, descreve-se a implementação dos modelos de aprendizagem.

Capítulo 5

Implementação dos Modelos de Aprendizagem

Este capítulo integra a implementação dos modelos de aprendizagem definindo o processo de aquisição do conhecimento, bases de treinamento, teste e regras, juntamente com as variáveis de entrada e saída com seus respectivos CDs. Faz-se uma descrição da arquitetura do SisND/EPs e da metodologia de avaliação dos modelos de aprendizagem avaliados.

5.1 Considerações Iniciais

Os modelos propostos integram os dois paradigmas na etapa de representação do conhecimento (RNAs e LD) com o objetivo de sanar problemas inerentes a um ou outro paradigma. Existem aplicações em que se faz necessária a utilização de formalismos de representação do conhecimento mais flexíveis, razão por que são avaliados vários modelos de aprendizagem envolvendo os SNDs.

O interesse em investigar diferentes operações aritméticas difusas baseia-se na premissa de que o processo de aprendizagem de máquina nos SNDs é caracterizado, além da interpretação das regras, pela forma como as regras conseguem emular o processo do pensamento humano. Dessa forma, é necessário investigar o resultado obtido implementando diferentes operações aritméticas difusas.

5.2 Implementação da Base de Conhecimento

Nesta tese, para efeitos de comparação, foram implementadas duas arquiteturas de SNDs utilizando os mesmos parâmetros de configuração. A primeira arquitetura, NEFCLASS, implementou as operações aritméticas difusas padrões *max/min* (Nauck e Kruse, 1995, 1996). A segunda, denominada de SisND/EPs, utilizou o algoritmo de retropropagação modificado e, além de implementar as operações aritméticas difusas padrões *max/min*, implementou as operações aritméticas difusas *t-normas* (Produto Algébrico, Produto de Hamacher, Produto de Einstein) e as *t-conormas* (Soma Algébrica, Soma de Hamacher, Soma de Einstein) na composição dos modelos

de neurônios difusos, juntamente com os operadores lógicos $OU(E)$ na camada intermediária e de saída da RNA, respectivamente.

Na seqüência, faz-se uma definição dos parâmetros de configuração utilizados por ambas as arquiteturas.

5.2.1 Aquisição e Representação do Conhecimento

Segundo Forsythe (1989), uma das etapas da aquisição do conhecimento está relacionada com a coleta das informações a partir de um ou mais especialistas ou através de outras fontes de conhecimento, por exemplo, livros e documentos até a sua codificação de forma computadorizada. A elicitación do conhecimento diz respeito às várias técnicas utilizadas na etapa de aquisição do conhecimento, como, entrevistas e a técnica de *teachback*.

Nesta tese, na etapa de elicitación do conhecimento do especialista de domínio, foram utilizadas as técnicas de análise de textos na bibliografia especializada da área da neurologia, entrevista direcionada e não-direcionada com o especialista e a técnica da observação no desempenho das suas tarefas. As entrevistas com os pacientes foram individuais ou com a presença de uma testemunha, com duração média de 1 hora. A escolha por essas técnicas deve-se ao fato de serem as que melhor se enquadraram no objetivo proposto. Após esta etapa, foram definidas formas de representar o conhecimento. Inicialmente, o conhecimento foi representado por meio de grafos do conhecimento. A explicitação do conhecimento do especialista utilizando grafos do conhecimento, mostrou-se adequada, devido a facilidade de visualizar o conhecimento a partir do especialista e da facilidade de compreensão da área de domínio²³. A seguir, foram extraídas as regras e exemplos do especialista de domínio. As regras consistem de uma parte *SE* que indica o antecedente, descrevendo os sintomas da doença; a parte *ENTÃO* expressa o conseqüente, descrevendo o diagnóstico final (*EE* ou *ENE*).

O tipo de abordagem utilizada envolveu os dados quantitativos relacionados com as variáveis discretas (quantidade sintomas apresentados) e, após a *fuzzificação*, dados qualitativos, relacionados com as variáveis contínuas. Cada uma dessas regras apresentou um grau de pertinência, correspondendo ao valor que foi colocado nas entradas da RNA.

A amostra utilizada na composição das bases de treinamento e de teste foi caracterizada como não aleatória, utilizando seleção de conveniência.

23

Os grafos do conhecimento encontram-se descritos no Apêndice 5. A metodologia para sua obtenção foi validada a partir de entrevistas com o médico neurologista que acompanhou o desenvolvimento da pesquisa.

5.2.2 Base de Regras

Foi gerada uma base inicial contendo 59 regras, referente aos sintomas apresentados pelos pacientes, de acordo com as variáveis de entrada do sistema. Após a implementação do SND, iniciou-se a fase de treinamento da RNA e, de acordo com as variáveis de entrada, ou seja, sintomas apresentados pelos pacientes, a base de regras foi sendo alterada, de acordo com os sintomas apresentados pelos pacientes. Atualmente o SND está com 149 regras, as quais representam, até o momento, os sintomas apresentados pelos pacientes. No tratamento das regras foi utilizado o SID *Mamdani*.

Para exemplificar, faz-se uma descrição de duas regras implementadas no SisND/EPs, utilizando as operações aritméticas difusas padrões *min*. Os graus de pertinência aparecem na parte antecedente e na parte conseqüente.

· Se $x_1 = \text{Às_vezes} (0, 3)$ e $x_2 = \text{Às_vezes} (0, 2)$ e $x_3 = \text{Às_vezes} (0, 4)$ e $x_4 = \text{Às_vezes} (0, 5)$ e $x_5 = \text{Às_vezes} (0, 3)$ e

$x_6 = \text{Sempre} (0, 6)$ e $x_7 = \text{Às_vezes} (0, 4)$ e $x_8 = \text{Às_vezes} (0, 3)$ e $x_9 = \text{Às_vezes} (0, 6)$ e

$x_{10} = \text{Freqüentemente} (0, 8)$ e $x_{11} = \text{Sempre} (0, 8)$ e $x_{12} = \text{Pequeno} (0, 2)$ e $x_{13} = \text{Sempre} (0, 5)$ e $x_{14} = \text{Freqüentemente} (0, 4)$

ENTÃO $y = EE (0, 2)$;

· Se $x_1 = \text{Sempre} (0, 2)$ e $x_2 = \text{Sempre} (0, 5)$ e $x_3 = \text{Sempre} (0, 4)$ e $x_4 = \text{Às_vezes} (0, 1)$ e $x_5 = \text{Às_vezes} (0, 4)$ e

$x_6 = \text{Sempre} (0, 7)$ e $x_7 = \text{Às_vezes} (0, 3)$ e $x_8 = \text{Às_vezes} (0, 4)$ e $x_9 = \text{Às_vezes} (0, 8)$ e $x_{10} = \text{Às_vezes} (0, 3)$ e

$x_{11} = \text{Às_vezes} (0, 4)$ e $x_{12} = \text{Grande} (0, 5)$ e $x_{13} = \text{Às_vezes} (0, 4)$ e $x_{14} = \text{Às_vezes} (0, 3)$

ENTÃO $y = ENE (0, 1)$.

onde:

$x_1, x_2, \dots, x_{14} =$ variáveis *fuzziificadas* de entrada que compõem os termos antecedentes das regras²⁴;

Às_vezes, freqüentemente, sempre = conjunto de valores lingüísticos;

$y =$ variável de saída do sistema que define um *EE* ou um *ENE*.

A etapa de construção da base de regras envolvendo os *EEs* e os *ENEs* foi realizada na cidade de Passo Fundo com a orientação do médico neurologista Hugo José Teixeira de Carvalho Filho.

²⁴ Os formulários de entrada contendo essas variáveis encontram-se no Anexo 1 (Formulário para coleta de dados do diagnóstico em eventos não epiléticos) e no Anexo 2 (Coleta de dados do diagnóstico em eventos epiléticos).

5.2.3 Bases de Treinamento e de Teste

Nesta tese, a base de casos é subdividida em uma base de treinamento e uma base de teste. Ambas as bases contêm 122 registros, caracterizando uma análise de 1 para 1²⁵. O processo de treinamento da RNA é interrompido quando o sistema alcança o menor erro q (EMQ)²⁶. Foi escolhido o critério de erro fixo porque se optou por não utilizar um conjunto de validação, em função do número de exemplos que integram a base de casos.

Para a implementação da base de casos foram cadastrados 244 pacientes que apresentavam EPs. Do total de casos selecionados, 105 apresentaram *EEs* e 139, *ENEs*. Dos 244 casos analisados, 122 foram escolhidos aleatoriamente para a composição da base de treinamento, ou seja, 52 pacientes com *EEs* e 70 pacientes com *ENEs*. Os demais casos foram utilizados na implementação da base de testes totalizando 53 pacientes com *EEs* e 69 pacientes com *ENEs*.

5.2.4 Variáveis de Entrada

Após uma análise nas principais características que envolvem os EPs, realizada em conjunto com o médico neurologista, foram escolhidas, dentre as variáveis que envolvem a definição de *EEs* e *ENEs*, 14 características presentes nos EPs de acordo com ILAE (1981), Costa et al. (1998), Luther et. al. (1982) e Freud (1998). As variáveis de entrada foram representadas utilizando números difusos²⁷.

As características que acometem os *ENEs* são definidas como:

(1) Quadros clássicos de conversão; (2) Distúrbios de sensibilidade e da atividade sensorial; (3) Fatores emocionais; (4) Transtornos; (5) Síncopes e (6) Distúrbios. As características que acometem os *EEs* são classificadas em (7) Sinais e sintomas parciais; (8) Durante o desmaio viu alguma coisa; (9) Durante o desmaio ouviu alguma coisa; (10) Durante o desmaio mordeu a língua; (11) Durante o desmaio machucou-se; (13) Eventos pós-críticos e (14) Sinais e sintomas generalizados²⁸.

As funções de pertinência foram definidas em conjunto com o médico especialista levando em consideração os sintomas e o seu grau de confiança em relação ao diagnóstico de *EEs* e de *ENEs*. Na prática sabe-se que é recomendável que as funções de pertinência sejam definidas em conjunto com

²⁵ A descrição da base de casos encontra-se nos Apêndices 1 e 2.

²⁶ O critério de parada baseado no valor mínimo do EMQ foi sugerido por Haykin (2001).

²⁷ Uma descrição das variáveis de entrada encontra-se no Apêndice 3.

²⁸ Para uma descrição mais detalhada do significado dos sintomas, consultar o Apêndice 6.

o médico especialista, com base na sua experiência. As funções escolhidas representaram de forma satisfatória as informações fornecidas pelo especialista.

As variáveis de entrada possuem as seguintes descrições:

Universo de Discurso: [0,100]

Domínio da Variável: (Percentual atingido)

Valores Lingüísticos/Função de Pertinência/Universo de Discurso:

- Às_Vezes - *Z* - [6, 45]
- Freqüentemente - *PI – SINO* - [6, 45, 54, 95]
- Sempre - *S* - [54, 95]

A Figura 5.1.1 ilustra o conjunto difuso das variáveis acima descritas.

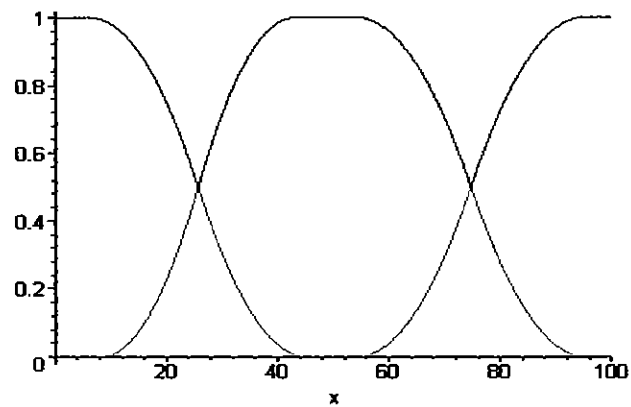


Figura 5.1.1 - Conjunto difuso das variáveis lingüísticas

As expressões analíticas das variáveis acima são definidas de acordo com os seguintes parâmetros:

· <nome da variável > = Às_Vezes: (Função *Z*)

$$\begin{aligned}
 &1 && \text{se } x < 6 \\
 &1 - 2((x - 6)/(45 - 6))^2 && \text{se } 6 \leq x \leq (6 + 45)/2 \\
 &2((x - 45)/(45 - 6))^2 && \text{se } (6 + 45)/2 < x \leq 45 \\
 &0 && \text{se } x > 45
 \end{aligned}$$

· <nome da variável > = Freqüentemente: (Função *PI – SINO*)

$$\begin{aligned}
 &0 && \text{se } x < 6 \\
 &2((x - 6)/(45 - 6))^2 && \text{se } 6 \leq x \leq (6 + 45)/2 \\
 &1 - 2((x - 45)/(45 - 6))^2 && \text{se } (6 + 45)/2 < x \leq 45 \\
 &1 && \text{se } 45 < x < 54
 \end{aligned}$$

$$\begin{array}{lll}
 1 - 2((x - 54)/(95 - 54))^2 & \text{se} & 54 \leq x \leq (54 + 95)/2 \\
 2((x - 95)/(95 - 54))^2 & \text{se} & (54 + 95)/2 < x \leq 95 \\
 0 & \text{se} & x > 95
 \end{array}$$

· <nome da variável> = Sempre: (Função S)

$$\begin{array}{lll}
 0 & \text{se} & x < 54 \\
 2((x - 54)/(95 - 54))^2 & \text{se} & 54 \leq x \leq (54 + 95)/2 \\
 1 - 2((x - 95)/(95 - 54))^2 & \text{se} & (54 + 95)/2 < x \leq 95 \\
 1 & \text{se} & x > 95
 \end{array}$$

A variável (12) Tempo de duração da crise relaciona-se ao tempo de duração de um EE, apresentando as seguintes configurações:

Universo de Discurso: [0,120]

Domínio da Variável: (Duração em minutos)

Valores Lingüísticos/Função de Pertinência/Universo de Discurso:

- Pequeno - Z - [1,2]
- Médio - PI - SINO - [1, 2, 4, 6]
- Grande - S - [4, 6]

A Figura 5.1.2 ilustra o conjunto difuso da variável descrita acima.

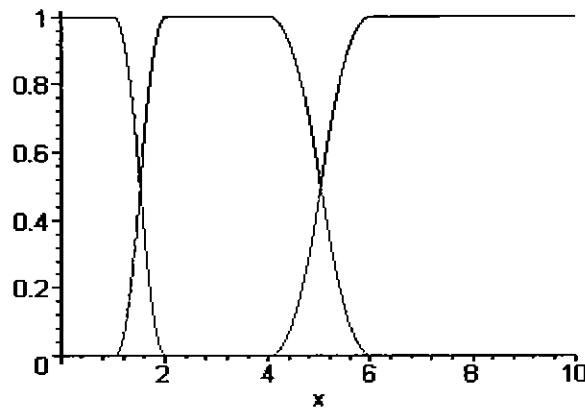


Figura 5.1.2 - Variável lingüística "tempo de duração"

As expressões analíticas da variável tempo de duração são definidas de acordo com os seguintes parâmetros:

· <nome da variável> = Pequeno: (Função Z)

$$\begin{array}{lll}
 1 & \text{se} & x < 1
 \end{array}$$

$$\begin{array}{ll}
 1 - 2((x - 1)/(2 - 1))^2 & \text{se } 1 \leq x \leq (1 + 2)/2 \\
 2((x - 2)/(2 - 1))^2 & \text{se } (1 + 2)/2 < x \leq 2 \\
 0 & \text{se } x > 2
 \end{array}$$

· <nome da variável > = Médio: (Função PI - SINO)

$$\begin{array}{ll}
 0 & \text{se } x < 1 \\
 2((x - 1)/(2 - 1))^2 & \text{se } 1 \leq x \leq (1 + 2)/2 \\
 1 - 2((x - 2)/(2 - 1))^2 & \text{se } (1 + 2)/2 < x \leq 2 \\
 1 & \text{se } 2 < x \leq 4 \\
 1 - 2((x - 4)/(6 - 4))^2 & \text{se } 4 < x \leq (5 + 5)/2 \\
 2((x - 6)/(6 - 4))^2 & \text{se } (5 + 5)/2 < x \leq 6 \\
 0 & \text{se } x > 6
 \end{array}$$

· <nome da variável > = Grande: (Função S)

$$\begin{array}{ll}
 0 & \text{se } x < 4 \\
 2((x - 4)/(6 - 4))^2 & \text{se } 4 \leq x \leq (4 + 6)/2 \\
 1 - 2((x - 6)/(6 - 4))^2 & \text{se } (4 + 6)/2 < x \leq 6 \\
 1 & \text{se } x > 6
 \end{array}$$

Frequência da Crise

Esta variável lingüística não é *fuzzificada*, ou seja, constitui uma variável numérica e está relacionada à frequência com que o paciente apresentou um *EE* ou um *ENE*. A partir desta variável, são definidos os respectivos universos de discursos das variáveis lingüísticas descritas.

Universo de Discurso: [0,30]

Domínio da variável: (Frequência das crises).

Para a definição dos valores dos CDs, os quais representam o universo de discurso, foi realizada a técnica "gerar e testar", ou seja, os valores foram gerados e testados. A seguir, os resultados foram comparados com os sintomas apresentados pelos pacientes. O objetivo foi concentrado em encontrar o número suficiente de regras para classificar corretamente o conjunto de padrões de entrada.

A descrição detalhada das variáveis de entrada encontra-se no Apêndice 5.

5.2.5 Variável de Saída

A variável de saída corresponde a um *EE* ou *ENE*. É importante ressaltar que partes do formulário de coleta de dados do diagnóstico em epilepsia foram retiradas do texto de Fernandes (1992, 1993), cujo questionário foi validado num estudo populacional na cidade de Porto Alegre/RS.

Na próxima seção, faz-se uma descrição da arquitetura do SisND/EPs proposta.

5.3 Arquitetura do SisND/EPs

Na implementação do SisND/EPs foi utilizado o SID *Mamdani* (Mamdani, 1974), baseado no MLP e no algoritmo de retropropagação modificado²⁹, o qual é utilizado na classificação de padrões e na geração das regras. Optou-se pelo SID *Mamdani* porque a arquitetura NEFCLASS utilizou este SID e, uma vez que os resultados de ambos os modelos foram comparados, devem utilizar a mesma forma de inferência.

5.3.1 Modelagem do Ambiente

Foi utilizada uma arquitetura neuro-difusa estática, descrita através de cinco camadas. A arquitetura foi baseada nos modelos propostos por Mitra e Pal (1992) e Pedrycz (1991, 1993). No entanto, foram realizadas algumas modificações no modelo, ou seja, Mitra e Pal utilizaram funções de pertinência como números difusos "pi". Nesta tese foram utilizadas funções "z", "pi-sino" e "s" porque representaram melhores os modelos. No protótipo do SisND/EP utilizou-se a linguagem de programação Delphi 7.0 na implementação da base de conhecimento híbrida. O algoritmo incorporou os pesos sinápticos utilizando os operadores lógicos *E/OU* na camada intermediária e de saída, respectivamente, no lugar da soma dos pesos.

As operações produto escalar, entre as entradas da RNA e seus pesos, foram substituídas pela operação *T*, definida pelas operações aritméticas difusas *t-normas*: Produto Algébrico; Produto de Hamacher; Produto de Einstein e pela operação aritmética difusa padrão *min*.

A operação somatório foi substituída pela operação *S*, a qual representa as operações aritméticas difusas *t-conormas*: Soma Algébrica; Soma de Hamacher; Soma de Einstein e pela operação aritmética difusa padrão *max*.

A função de ativação adotada na implementação dos modelos foi a função logística. Foi escolhida esta função porque os valores de saída foram definidos no intervalo [0,1]. A função é representada pela Equação:

$$\varphi(z) = \frac{1}{1 + e^{-kz}} \quad (5.3.1)$$

onde:

k = parâmetro de inclinação da função logística.

²⁹ Inspirado nos trabalhos de Rumelhart et al., (1986).

A derivada da função logística é dada através da Equação:

$$\varphi'(z) = \varphi(z) * [1 - \varphi(z)] \tag{5.3.2}$$

As entradas da RNA apresentam valores quantitativos (numéricos). A saída da RNA consiste em determinar, a partir dos valores de entrada, o tipo de EP (*EE* ou um *ENE*) apresentado pelos pacientes. A RNA foi treinada a partir de uma base de treinamento, generalizando-a na fase de testes com exemplos desconhecidos e que não foram treinados.

Na implementação dos modelos de aprendizagem utilizou-se uma RNA com 14 unidades de entrada. Para cada unidade de entrada u_j existem três valores lingüísticos "às_vezes", "freqüentemente" e "sempre".

A modelagem do ambiente foi composta por entradas, objetivos e pesos difusos integrando um SND Tipo 3. De acordo com Fullér (2000), um SND tipo 3 é utilizado na implementação de regras *SE-ENTÃO*. Os pesos são atualizados através do método do gradiente descendente (Haykin, 2001), com exceção das operações aritméticas difusas padrões *max/min*, as quais fazem uso da função sinal do gradiente.

A Figura 5.1.3 exemplifica a arquitetura do SisND/EPs, a qual possui neurônios *E/OU* na camada intermediária e de saída, respectivamente.

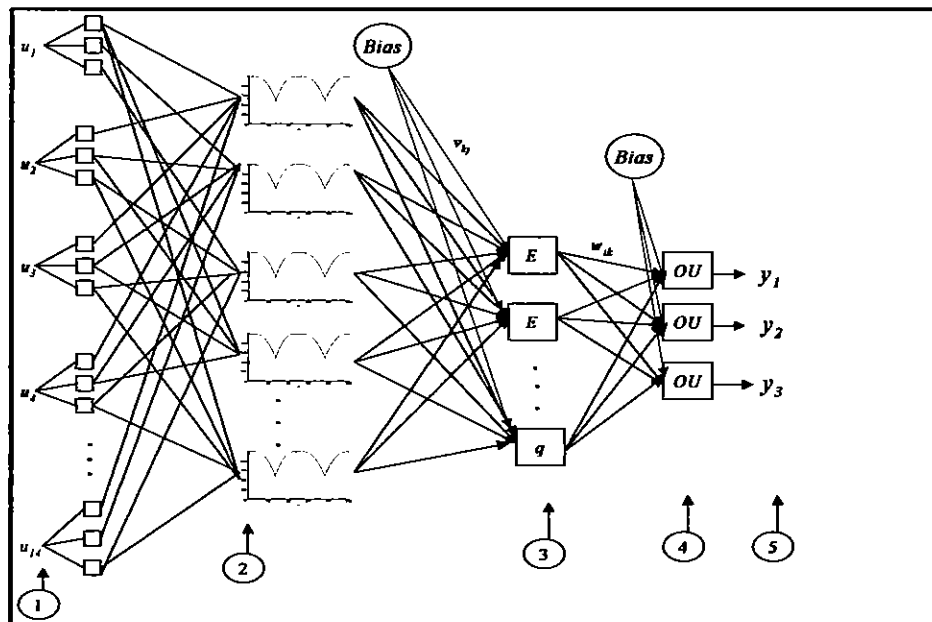


Figura 5.1.3 - Arquitetura do SisND/EPs

A seguir, faz-se a descrição das camadas que compõem o SisND/EPs:

· CAMADA 1 (*Entrada*)

Cada nó da camada 1 representa uma variável de entrada da RNA. É a camada que recebe os dados a partir do ambiente externo, conforme a Equação:

$$u_j^p = (u_1^p, \dots, u_j^p, \dots, u_n^p) \quad (5.3.3)$$

onde:

u_j^p = valores de entrada j para o padrão p .

A camada 1 é responsável pelo pré-processamento das informações de entrada, ou seja, as informações são transmitidas para a camada 2 baseadas no valor da percentagem obtida, utilizando o número total de crises que o paciente apresentou em relação às demais variáveis de entrada da RNA.

· CAMADA 2 (*Graus de Pertinência*):

Na camada 2 são calculados os graus de pertinência a partir dos valores de entrada $crisp (u_1^p, \dots, u_j^p, \dots, u_n^p)$:

$$x_j^p = (x_1^p, \dots, x_j^p, \dots, x_n^p) \in [0, 1] \quad (5.3.4)$$

onde:

x_j^p = valor de entrada *fuzziificado* j para o padrão p .

Esta camada fuzzifica os valores de entrada verificando em que conjuntos os valores se incluem. Nesta pesquisa foram definidos três conjuntos, compostos pelas variáveis linguísticas "às_vezes", "frequentemente" e "sempre".

· CAMADA 3 (*neurônio difuso E*):

Nesta camada foram implementadas as regras de classificação utilizando o neurônio difuso E a partir das operações aritméticas-difusas:

- Produto Algébrico;
- Produto de Hamacher;
- Produto de Einstein;
- Operação aritmética difusa padrão *min.*

Cada neurônio da terceira camada está conectado a todos os neurônios da quarta camada.

O número de neurônios da camada intermediária foi de 149 neurônios em todos os modelos implementados, representando 149 regras difusas.

· CAMADA 4 (*neurônio difuso OU*):

Cada nó da camada 4 representa o cálculo do neurônio difuso OU utilizando as operações aritméticas difusas:

- Soma Algébrica;
- Soma de Hamacher;
- Soma de Einstein;
- Operação aritmética difusa padrão *max*.

Nesta camada os valores *fuzzi ficados* que chegam da camada anterior servem de entrada para as próximas funções, constituída pela respectiva operação aritmética difusa *Soma*.

- CAMADA 5 (*Saída*):

Esta é a camada de saída do SisND/EPs (y_1, y_2, y_3), a qual poderá representar um *EE* (saída binária 100) ou um *ENE* (saída binária 010).

A Figura 5.1.3 representa uma RNA composta de três camadas. O que difere é a maneira de referenciar a RNA. Neste modelo foi definida a camada 1, como sendo a camada apenas de entrada e, a camada 5, como saída. Esta opção de RNA foi escolhida tendo em vista a necessidade de representar as operações aritméticas difusas *t-normas* e *t-conormas*, juntamente com os operadores lógicos OU(E).

5.3.2 Interface

A interface do SisND/EPs, encarregada da comunicação entre o sistema e o mundo exterior, compõe-se de duas fases:

(*Fase_1*): utilizou tabelas para facilitar a conferência;

(*Fase_2*): o processamento é realizado utilizando matrizes para possibilitar maior velocidade no treinamento da RNA.

A implementação por meio de tabela compõem-se de quatro módulos:

- (1) parâmetros de configuração;
- (2) inclusão/exclusão/alteração das regras;
- (3) fase de conferência do SisND/EPs;
- (4) tela de ajuda ao usuário.

A partir do *menu* principal o usuário tem acesso aos arquivos do sistema: bases de treinamento, teste, regras e ao cadastro dos pacientes.

A seguir é apresentada a interface do sistema.

No primeiro módulo, o usuário define os parâmetros: valor do *bias* da camada intermediária e de saída (θ_k, θ_i); peso do *bias* da camada intermediária e de saída (v_{k0} e w_{i0}); taxa de aprendizagem,

critério de convergência, valor do termo "momento" e o número de épocas, como mostra a Figura 5.1.4.

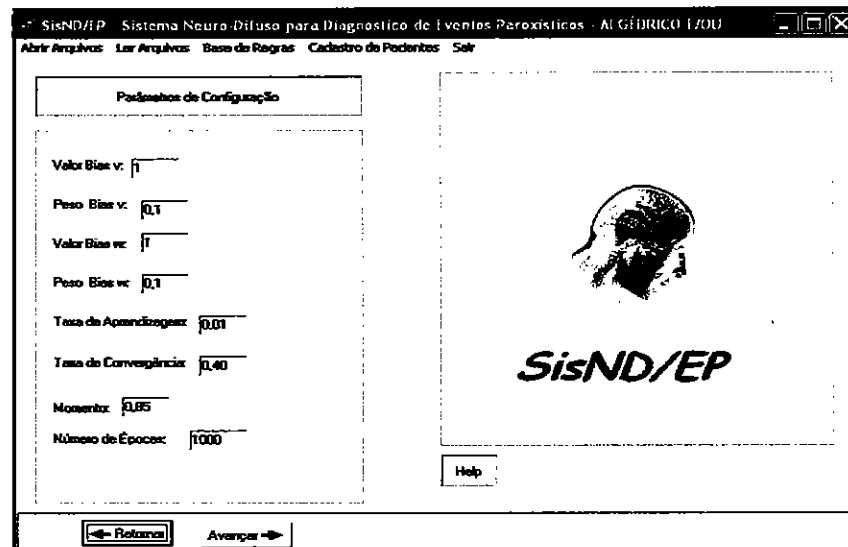


Figura 5.1.4 - Parâmetros de configuração

Na fase de treinamento foram utilizados os valores:

(a) valor e peso do termo *bias* positivo: 1 e 0,1, respectivamente; (b) taxa de aprendizagem de 0,01; (c) termo *momento* de 0,85; (d) pesos sinápticos aleatórios no intervalo [-0,5 até +0,5].

O critério de parada na fase de treinamento foi baseado no menor EMQ e no número de épocas (Haykin, 2001). O número de épocas realizado por cada modelo de estrutura foi de 1.000 iterações. Foi escolhido este número de épocas tendo em vista que, a partir de 1.000 épocas, o EQM permaneceu instável.

Na avaliação da aprendizagem são utilizados: (a) valor e peso do termo *bias* negativos; (b) termo "momento" de 0,95; (c) pesos fixos em 0,01³⁰.

O SisND/EPs possui um módulo que permite o cadastro dos pacientes, no qual os dados são digitados diretamente a partir dos Anexos 1 e 2. Por meio do menu Abrir Arquivos são inseridos registros nas bases de treinamento e de teste. Os próximos parâmetros a serem configurados fazem referência aos neurônios difusos propostos por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994).

A seguir é selecionada a operação aritmética difusa:

- (a) Soma/Produto Algébrico OU(E);
- (b) Soma/Produto de Hamacher OU(E);

(c) Soma/Produto de Einstein OU(E);

(d) *max/min*.

Estas operações aritméticas difusas, juntamente com o tipo de neurônio difuso, são mostradas na Figura 5.1.5.

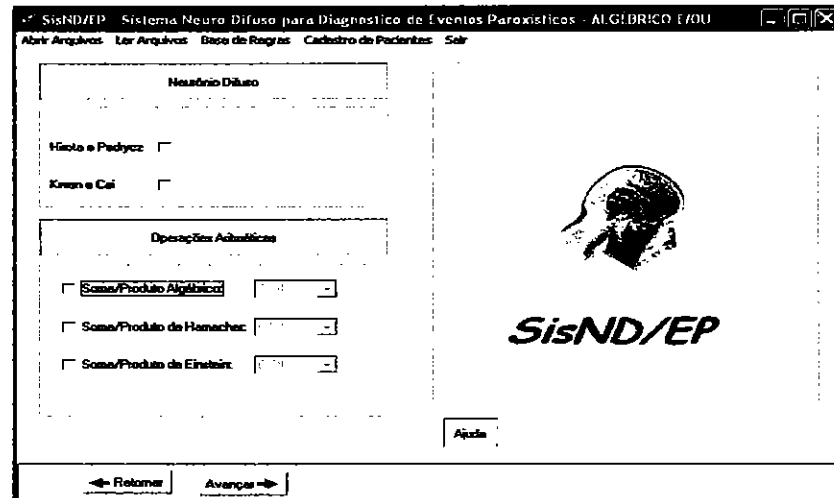


Figura 5.1.5 - Neurônios difusos e as operações aritméticas difusas

O segundo módulo implementou o processo de inclusão/exclusão e/ou alteração das regras, como mostra a Figura 5.1.6.

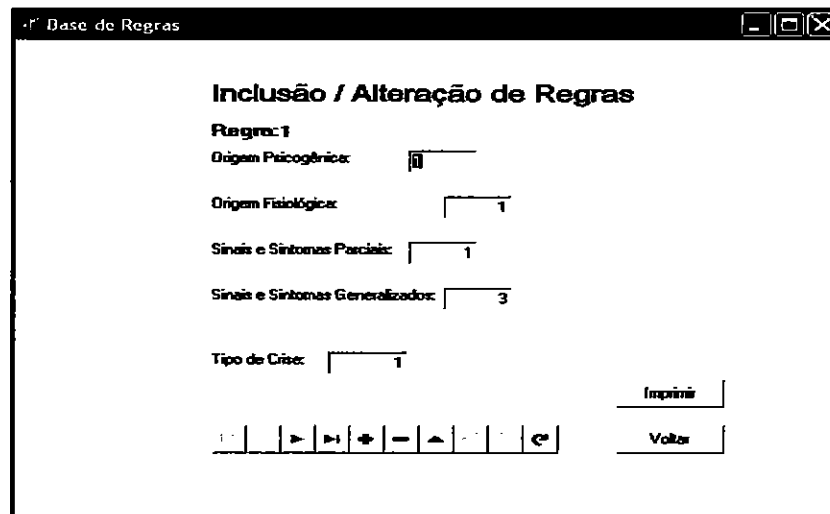


Figura 5.1.6 - Inclusão e alteração de regras

Neste módulo foram incluídas as regras que compõem o SisDN/EPs. O sistema foi inicializado com 59 regras.

Após ajustados os parâmetros de configuração, o terceiro módulo integrou a fase de conferência do SisND/EPs (Figura 5.1.7).

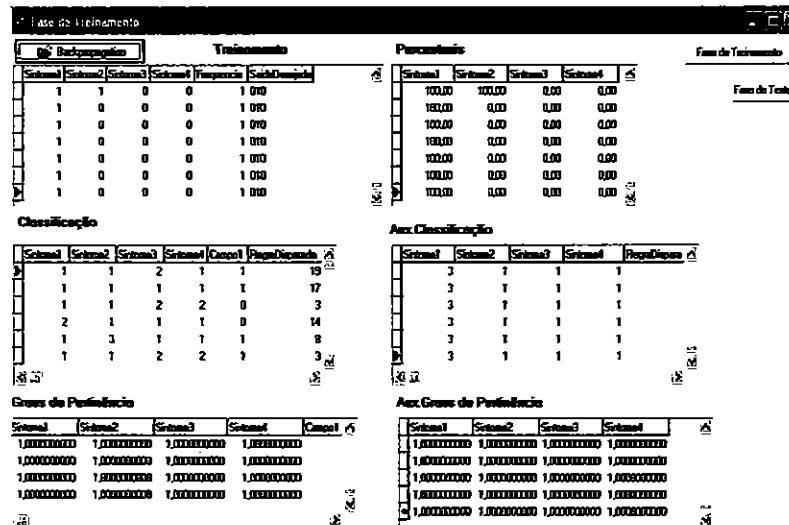


Figura 5.1.7 - Fase de conferência do SisND/EPs

De acordo com a Figura 5.1.7 o algoritmo implementado permite mostrar passo a passo as etapas envolvidas no processo de aprendizagem, incluindo:

- (a) fases de treinamento e de teste;
- (b) percentuais;
- (c) tabelas de classificação e de graus de pertinência.

Na tabela de classificação foram implementadas as entradas do sistema com seus respectivos valores lingüísticos "às_vezes", "freqüentemente" e "sempre". A tabela dos graus de pertinência mostra os valores de entrada *fuzziificados*.

Optou-se por essa forma de algoritmo porque, à medida que o programa foi sendo executado, os resultados foram conferidos, evitando e/ou corrigindo eventuais erros que poderiam surgir a partir do treinamento da RNA.

O quarto módulo integra a tela de ajuda ao usuário, descrevendo as principais características relacionadas com os EPs, entre elas:

- (a) classificação e características dos *EEs* e
- (b) classificação e características dos *ENEs*.

A Figura 5.1.8 mostra a tela de ajuda ao usuário.

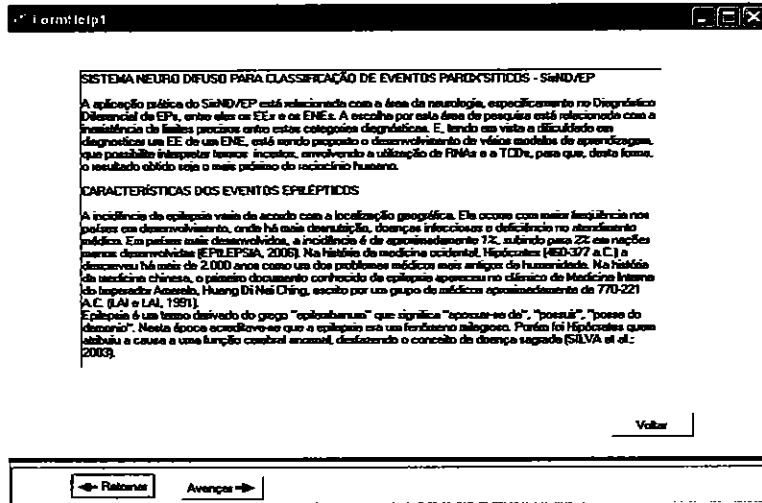


Figura 5.1.8 - Tela de ajuda ao usuário

A implementação utilizando matrizes compõe-se de dois módulos:

- (1) fase de treinamento e de teste;
- (2) visualização do processo de treinamento da RNA.

No primeiro módulo, composto pela fase de treinamento e de teste, são apresentados os resultados obtidos pela RNA: saída desejada, saída real, EMQ, faixa de EMQ, juntamente com os parâmetros de configuração valor e peso do bias, taxa de aprendizagem, termo momento, faixa de pesos (Figura 5.1.9).

TREINAMENTO			
	Percentual	Classificação dos Registros	Classificação Analisa
1	0.000000000000000000	01111111111122101	0111211111112132200
2	0.000000000000000000	01111111111122101	011121112111132100
3	0.000000000000000000	01111111111122101	01111111111113200
4	0.000000000000000000	01111111111122101	01111111111113100
5	0.000000000000000000	01111111111122101	01111111111113100
6	0.000000000000000000	01111111111122101	01111111111113100
7	0.000000000000000000	01111111111122101	01111111111113100
8	0.000000000000000000	01111111111122101	01111111111113100
9	0.000000000000000000	01111111111122101	01111111111113100
10	0.000000000000000000	01111111111122101	01111111111113100
Resumo:			
103 ENE	43 EE	01102071111102071002050307	0111011111110102050
120 ENE	49 EE	0110206111102061111002051	0110206111102061111002051
121 ENE	105 EE	01111111111100460325	01111111111100460325
122 ENE	171 EE	01046039008310390111711002031	01046039008310390111711002031
Total TCC			
EE: 47	EE: 9	011110086111100861100821	0111110086111100861100821
ENE: 70	ENE: 0	01023026402731111111102071	01023026402731111111102071
Resumo de Última época:			
100 1 época 1 0 0		01023026402731111111102071	
100 1 Saída	0.000000000000000000	010230264027311111102071	010230264027311111102071
100 1 Erro (MSE)	0.000000000000000000	010230264027311111102071	010230264027311111102071
ERRO TOTAL 0.211372963452 CONVERGIDO			
100 2 época 1 0 0		010230264027311111102071	
100 2 Saída	0.000000000000000000	010230264027311111102071	010230264027311111102071
100 2 Erro (MSE)	0.000000000000000000	010230264027311111102071	010230264027311111102071
ERRO TOTAL 0.000000000000 CONVERGIDO			
100 3 época 1 0 0		010230264027311111102071	
100 3 Saída	0.000000000000000000	010230264027311111102071	010230264027311111102071
100 3 Erro (MSE)	0.000000000000000000	010230264027311111102071	010230264027311111102071
ERRO TOTAL 0.000000000000 CONVERGIDO			
100 4 época 0 1 0		010230264027311111102071	
100 4 Saída	0.000000000000000000	010230264027311111102071	010230264027311111102071
100 4 Erro (MSE)	0.000000000000000000	010230264027311111102071	010230264027311111102071
ERRO TOTAL 0.000000000000 CONVERGIDO			

Figura 5.1.9 - Fase de treinamento e de teste

O segundo módulo mostra o processo de treinamento da RNA, ou seja, a saída RNA e a saída desejada podem ser acompanhadas no gráfico em movimento, época por época (Figura 5.1.10).

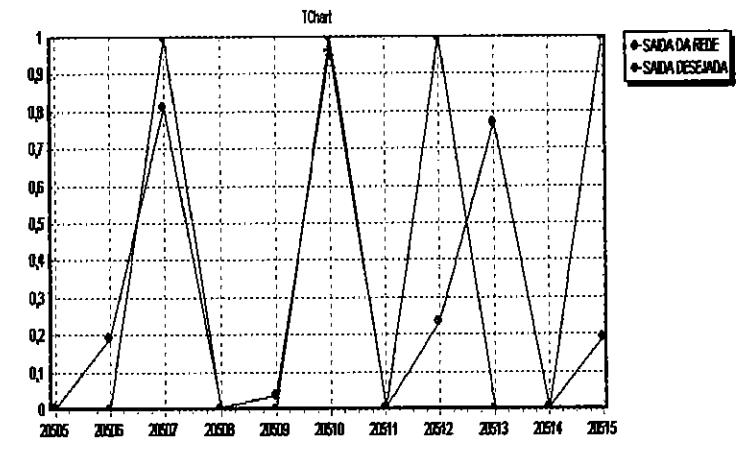


Figura 5.1.10 - Gráfico em movimento

De acordo com a Figura 5.1.10, pode-se conferir em qual iteração a RNA apresenta resultados de convergência ou não. O critério de convergência da RNA está relacionado com o erro quadrático total $E(t)$, conforme a Equação:

$$SE E(t) \leq 0.40 ENTÃO o registro convergiu \quad (5.3.5)$$

Este critério de convergência foi escolhido porque a arquitetura NEFCLASS utiliza o mesmo critério de convergência e os modelos avaliados são comparados com este modelo padrão.

No desenvolvimento do SisND/EPs foram utilizados os seguintes recursos de *software* e *hardware*: (a) ambiente de desenvolvimento Delphi 7.0; (b) microcomputador Athlon XP 2,8 Ghz, com memória de 512 MB. A escolha pela linguagem de programação deve-se às seguintes razões:

- possui componentes próprios para *Structured Query Language* (SQL), o que possibilita qualquer pesquisa e amostra dentro da base de dados;
- são disponibilizados na internet vários componentes extras que não fazem parte do pacote Delphi, como botões, o que facilita a programação, como por exemplo, o componente para criação de gráficos *chart*, armazenar valores em um *grid*, ajustes de precisão, etc;
- trabalha com índices, ponteiros, tabelas e uma grande variedade de Banco de Dados;
- tabelas e índices confiáveis e de fácil manipulação. Podem-se criar novos índices sem corromper os anteriores.

5.4 Avaliação dos Modelos de Aprendizagem

Segundo Nassar (1998), a avaliação de sistemas especialistas ocorre através de duas fases: (1) avaliação do desempenho do sistema e (2) validação interna da base de conhecimentos. Nesta tese, as análises estatísticas implementadas incluem a avaliação do desempenho dos modelos de aprendizagem e a validação interna da base de conhecimentos, ambas aplicadas na fase de teste do SisND/EPs.

Na seqüência a descrição das etapas de avaliação dos modelos.

5.4.1 Avaliação do Desempenho dos Modelos (Nassar, 1998)

Na avaliação do desempenho dos modelos optou-se pelo teste de hipótese³¹, o qual está fundamentado numa situação experimental. Foi escolhido o teste de hipótese para proporções tendo em vista o objetivo de comparar os resultados obtidos com as operações aritméticas difusas *t-normas* e *t-conormas* com as operações aritméticas difusas padrões já existentes, implementadas com a arquitetura NEFCLASS, aceitá-la ou rejeitá-la.

O teste de hipóteses para proporções (%) consiste na comparação do percentual de acertos da arquitetura do SisND/EPs (incluindo cada operação aritmética difusa *t-norma* e *t-conorma*), sendo esses resultados comparados com o percentual de acertos da arquitetura NEFCLASS através da implementação das operações aritméticas difusas padrões *max/min*, exemplificadas na Tabela 5.1.1.

Quadro 5.1.1 - Teste de hipóteses

Arquiteturas e Neurônio Difuso	Operação Aritmética	n	%acertos	P	Erro-padrão	z	p_valor
NEFCLASS	Max/Min						
SisND/EP - Hirota e Pedrycz	Soma/Produto Algébrico						
SisND/EP - Hirota e Pedrycz	Soma/Produto de Hamacher						
SisND/EP - Hirota e Pedrycz	Soma/Produto de Einstein						
SisND/EP - Hirota e Pedrycz	Max/Min						
SisND/EP - Kwan e Cai	Soma/Produto Algébrico						
SisND/EP - Kwan e Cai	Soma/Produto de Hamacher						
SisND/EP - Kwan e Cai	Soma/Produto de Einstein						
SisND/EP - Kwan e Cai	Max/Min						

onde:

n = total de casos da base de teste;

% *acertos* = percentual de acertos;

$$P = [(n_0 * p_0) + (n_1 * p_1)] / (n_0 + n_1);$$

onde:

³¹ Informações mais detalhadas com o teste de hipótese consultar Devore (2006).

p_0 = percentual de acertos das operações aritméticas difusas padrões *max/min*;

p_1 = percentual de acertos de cada uma das demais operações aritméticas difusas *t-normas* e *t-conormas*;

Erro-padrão = $\hat{\sigma}(p_0 - p_1)$;

z = expressa a distância entre o percentual de acertos das hipóteses alternativas;

p_valor = probabilidade de significância.

É utilizada a Equação:

$$z = \frac{p_1 - p_0}{\hat{\sigma}(p_0 - p_1)} \tag{5.4.1}$$

O desvio-padrão segue a Equação:

$$\hat{\sigma}(p_0 - p_1) = \sqrt{\frac{P(1 - P)}{n_0} + \frac{P(1 - P)}{n_1}} \tag{5.4.2}$$

onde:

n_0 e n_1 = total de casos da base de teste.

5.4.2 Validação Interna da Base de Conhecimentos (Nassar, 1998)

São adotadas as medidas de validade interna, exemplificadas no Quadro 5.1.2.

Quadro 5.1.2 - Medidas de validade interna dos SNDs

Diagnóstico do Sistema	Diagnóstico do Especialista Padrão Ouro		TOTAL
	Evento Epiléptico (Positivo)	Evento Epiléptico (Negativo)	
Evento Epiléptico (positivo)	N_{11}	falso positivo N_{12}	VP_P
Evento não Epiléptico (negativo)	falso negativo N_{21}	N_{22}	VP_n
TOTAL	SE_n	ES_P	n

O Quadro 5.1.2 define as medidas:

· Sensibilidade (SEn):

$$SEn = N_{11}/SEn \tag{5.4.3}$$

A SEn é definida como a probabilidade ou capacidade do sistema gerar resultado positivo, dado o caso ser positivo, ou seja, é a capacidade do sistema de detectar casos positivos.

· Especificidade (ESp):

$$ESp = N_{22}/ESp \quad (5.4.4)$$

A ESp é definida como a probabilidade de o sistema gerar um diagnóstico negativo, dado o caso ser negativo, ou seja, é a capacidade do sistema detectar casos negativos.

· Valor Preditivo Positivo (VPp):

$$VPp = N_{11}/VPp \quad (5.4.5)$$

O VPp é definido como a probabilidade do caso pertencer à hipótese diagnóstica em questão, dado o diagnóstico positivo do sistema. É útil quando se deseja saber o grau de certeza do resultado ser positivo quando o diagnóstico oferecido pelo sistema é positivo.

· Valor Preditivo Negativo (VPn):

$$VPn = N_{22}/VPn \quad (5.4.6)$$

O VPn é definido como a probabilidade do caso de ser negativo, dado um diagnóstico negativo do sistema.

Para qualificar o aprendizado da RNA foram definidas as medidas relacionadas com os percentuais de acertos: (1) adequado (acima de 80%, inclusive); (2) inadequado (abaixo de 80%). As medidas foram baseadas nas características de: (1) eficiência, ou seja, no desempenho apresentado pelas arquiteturas NEFCLASS e do SisND/EPs e (2) funcionalidade do sistema, se as arquiteturas geram resultados mais precisos por meio da validação interna de suas bases de conhecimento.

Essas medidas são descritas no próximo capítulo, após as fases de treinamento e de teste e na etapa de avaliação interna da base de conhecimentos.

5.5 Considerações Finais

Este capítulo descreveu a implementação dos modelos de aprendizagem, definindo as bases de treinamento e de teste, forma de aquisição do conhecimento, base de regras, juntamente com as variáveis de entrada e saída do sistema com seus respectivos CDs. Fez-se uma descrição do SisND/EP e da metodologia de avaliação dos modelos de aprendizagem propostos.

Os resultados obtidos a partir do SisND/EPs, o qual implementa as operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein

OU(E) e *max/min* foram comparados com os resultados da arquitetura NEFCLASS, que utiliza as operações aritméticas difusas padrões *max/min*.

Os resultados da investigação utilizando diferentes operações aritméticas difusas *t-normas* e *t-conormas* foram analisados considerando indicadores quantitativos de modelagem, os quais são descritos no próximo capítulo.

Capítulo 6

Resultados e Discussão

Neste capítulo são apresentados os resultados e discussões das investigações realizadas com as arquiteturas NEFCLASS e com o SisND/EPs incluindo os neurônios difusos propostos por Hirota-Pedrycz e Kwan-Cai.

6.1 Considerações Iniciais

Os resultados das investigações referentes aos modelos de aprendizagem foram analisados considerando os indicadores quantitativos de modelagem. Cada modelo foi avaliado a partir das mesmas bases de casos envolvendo os *EEs* e *ENEs*, segundo três eixos:

- (1) aprendizagem dos modelos;
- (2) avaliação de desempenho;
- (3) validação interna da base de conhecimentos.

Nas etapas de avaliação de desempenho e validação interna da base de conhecimentos são confrontados os percentuais de classificação obtidos por meio das operações aritméticas difusas com os percentuais de classificação fornecidos pelo especialista de domínio.

6.2 Aprendizagem dos Modelos

Na seqüência são apresentados os resultados obtidos na fase de treinamento utilizando as arquiteturas NEFCLASS e o SisND/EPs. Ambas as arquiteturas utilizaram uma base de casos composta por 122 pacientes, tanto na fase de treinamento como na fase de teste.

A arquitetura do SisND/EPs utiliza os neurônios difusos propostos por Hirota-Pedrycz (1992, 1994) e Kwan-Cai (1994) utilizando as operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein OU(E) e as operações aritméticas difusas padrões *max/min* (Quadro 3.1.1).

6.2.1 NEFCLASS

As operações aritméticas difusas padrões *max/min* implementadas com a arquitetura NEFCLASS apresentaram os resultados de EMQ por época visualizados na Figura 6.1.1.

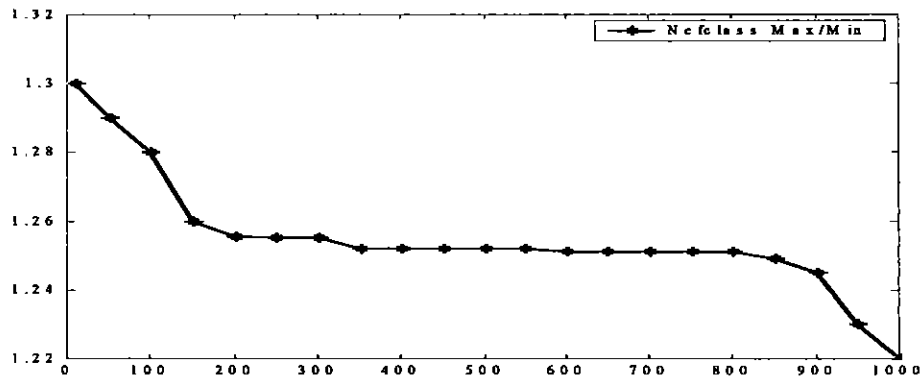


Figura 6.1.1 - EMQ - *Max/Min* para a arquitetura NEFCLASS

Como se pode observar na Figura 6.1.1, o EMQ obteve os maiores valores, permanecendo entre os valores [1, 30; 1, 22], apresentando uma média de EMQ de 1,22 e uma Variação da Taxa Absoluta do EMQ de 0,002. O percentual de acertos na fase de treinamento foi de 82,78%. Observou-se que o EMQ diminuiu até a época 200; após, a aprendizagem permaneceu estável e, somente a partir da época 900, a aprendizagem retornou a apresentar melhores resultados.

6.2.2 Hirota-Pedrycz

Após a implementação das operações aritméticas difusas Soma/Produto Algébrico OU(E), foram observados os seguintes resultados de EMQ por época (Figura 6.1.2).

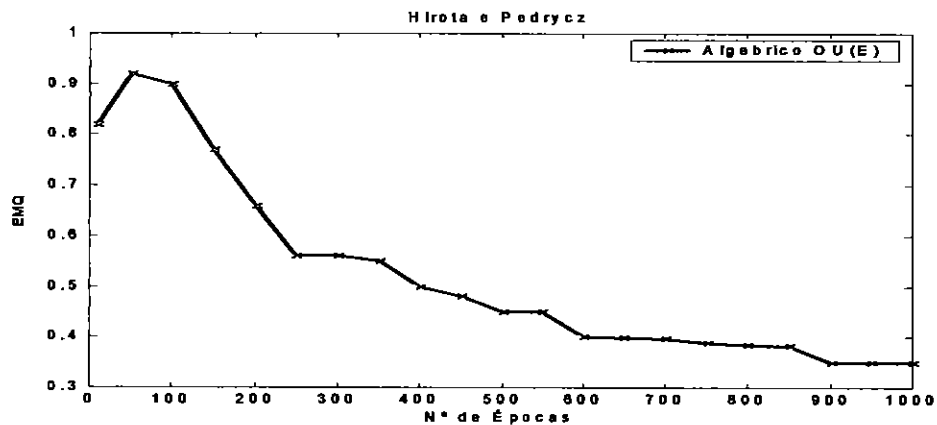


Figura 6.1.2 - EMQ - Soma/Produto Algébrico (OU/E) para o neurônio difuso Hirota-Pedrycz

Como se pode observar na Figura 6.1.2, o EMQ permaneceu entre os valores $[0,82; 0,35]$, apresentando uma média de EMQ de 0,51 e uma Variação da Taxa Absoluta do EMQ de 0,057. O percentual de acertos na fase de treinamento foi de 51,37%. Observou-se que o sistema convergiu até a época 900; após, a aprendizagem permaneceu estável.

As operações aritméticas difusas Soma/Produto de Hamacher OU(E) apresentaram os resultados de EMQ por época mostrados na Figura 6.1.3.

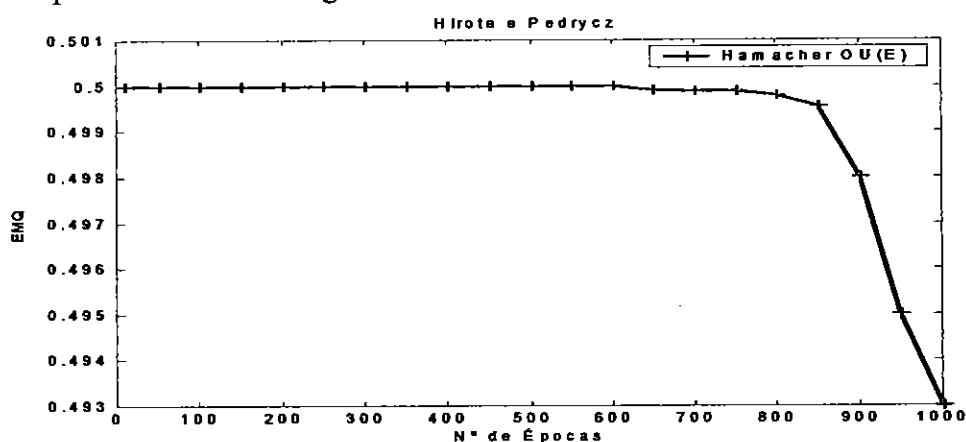


Figura 6.1.3 - RMS - Soma/Produto de Hamacher (OU/E) para o neurônio difuso Hirota-Pedrycz

De acordo com a Figura 6.1.3, observa-se que, até a época 800, o EMQ permaneceu estável em 0,499, apresentando baixos percentuais de acertos na fase de treinamento. Nesse ponto ocorreu uma paralisia na modificação dos pesos sinápticos. A partir da época 800, o EMQ diminuiu até 0,493.

As operações aritméticas difusas Soma/Produto de Einstein OU(E) apresentaram os resultados de EMQ por época observados na Figura 6.1.4.

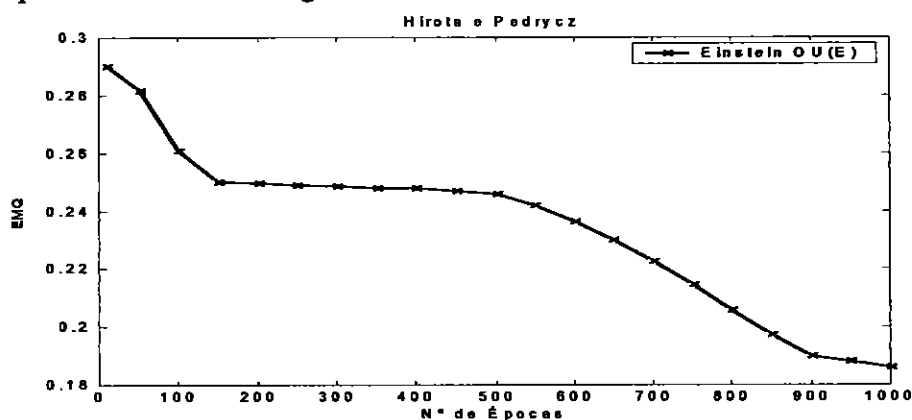


Figura 6.1.4 - EMQ - Soma/Produto de Einstein para o neurônio difuso Hirota-Pedrycz

De acordo com a Figura 6.1.4, o EMQ permaneceu entre os valores $[0, 29; 0, 18]$, com uma média de EMQ de 0,23 e apresentando uma Variação da Taxa Absoluta do EMQ em 0,002. O percentual de acertos na fase de treinamento foi de 87,70%. Observou-se que o sistema aprendeu até a época 190, após o que a aprendizagem permaneceu estável e, somente a partir da época 450, a aprendizagem retornou a convergir com a minimização do EMQ.

As operações aritméticas difusas padrões *max/min* apresentaram os resultados de EMQ por época mostrados na Figura 6.1.5.

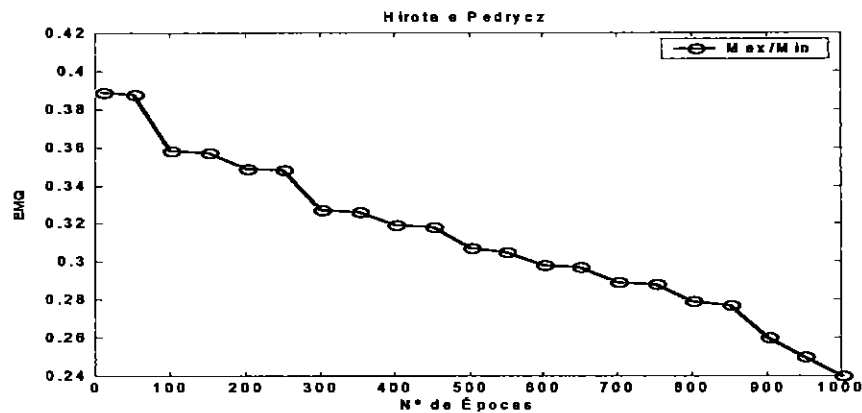


Figura 6.1.5 - EMQ - *Max/Min* para o neurônio difuso Hirota-Pedrycz

De acordo com a Figura 6.1.5, o EMQ permaneceu entre os valores $[0, 385; 0, 24]$, com uma média de EMQ de 0,24 e apresentando uma Variação da Taxa Absoluta do EMQ em 0,01. O percentual de acertos na fase de treinamento foi de 57,37%, porém a saída da RNA apresentou resultados iguais. Observou-se que a minimização do EMQ ocorreu de forma gradual próximo à época 880; a seguir, ocorreu o EMQ diminuiu. Os resultados obtidos na aprendizagem utilizando o neurônio difuso proposto por Hirota-Pedrycz são mostrados na Figura 6.1.6.

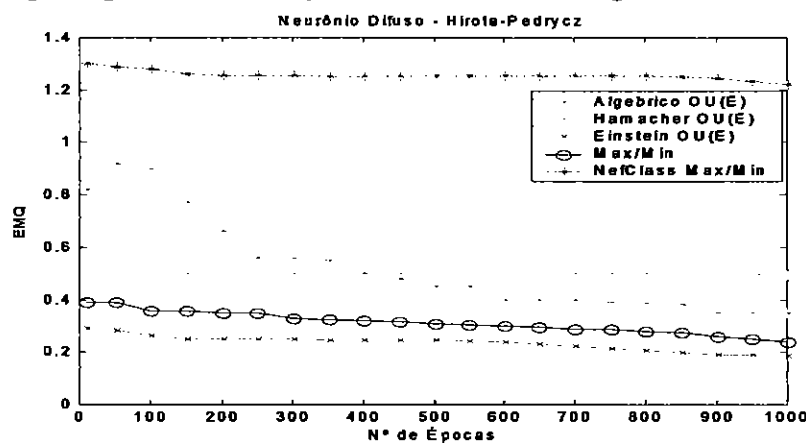


Figura 6.1.6 - EMQ da arquitetura NEFCLASS e do neurônio difuso Hirota-Pedrycz

De acordo com a Figura 6.1.6 os resultados mostram que as operações aritméticas difusas padrões *max/min*, implementadas com a arquitetura NEFCLASS, apresentaram o maior EMQ por época, ao passo que as operações aritméticas difusas Soma/Produto de Einstein OU(E), implementadas com o neurônio difuso proposto por Kwan-Cai, apresentaram o menor EMQ por época. As operações aritméticas difusas Soma/Produto Algébrico OU(E) apresentaram na fase inicial do treinamento uma subida do EMQ; após o EMQ diminuiu; os resultados obtidos com as operações aritméticas difusas Soma/Produto de Hamacher permaneceram constantes em todo o processo de aprendizagem. Observou-se que a aprendizagem a partir da arquitetura NEFCLASS, juntamente com as operações aritméticas difusas Soma/Produto Algébrico OU(E) e Soma/Produto de Einstein OU(E), aprendeu melhor em torno de 200 épocas de treinamento; após, a aprendizagem permaneceu estável próximo à época 500; a seguir, voltou a convergir.

Após a avaliação no processo de aprendizagem, foram excluídas da análise estatística as operações aritméticas difusas Soma/Produto de Hamacher OU(E) (Figura 6.1.3), por apresentarem problemas na fase de treinamento, entre eles: EMQ estável, ocorrendo paralisia na modificação dos pesos e problemas no critério de convergência. Foram excluídas também da análise estatística as operações aritméticas difusas padrões *max/min* (Figura 6.1.5), ambas implementadas com o neurônio difuso proposto por Hirota-Pedrycz, por apresentarem: saída da RNA com resultados iguais e instabilidade, ou seja, a RNA se manteve instável, mesmo os pesos sendo alterados, a RNA não classificou corretamente os padrões.

Algumas razões que levaram a RNA a apresentar estes resultados, incluem:

(a) após a fase de treinamento com as operações aritméticas difusas padrões *max/min*, especificamente em relação à implementação da função sinal do gradiente (E^p), correspondente ao erro no padrão p , verificou-se que o valor da variável β precisa ser "negativo", caso contrário, a saída da RNA fornece resultados iguais;

(b) analisando os resultados das operações aritméticas difusas Soma/Produto de Hamacher OU(E), o problema da paralisia na modificação dos pesos ocorreu em razão das unidades de saída da RNA conterem valores próximos de "0" ou "1".

O Quadro 6.1.1 descreve os resultados das implementações na fase de treinamento com as arquiteturas NEFCLASS e o SisND/EPs utilizando diferentes operações aritméticas difusas implementadas com o neurônio difuso proposto por Hirota-Pedrycz.

Os percentuais de acertos descritos referem-se aos acertos obtidos na fase de treinamento.

Quadro 6.1.1 - Desempenho das arquiteturas NEFCLASS e Hirota-Pedrycz: fase de treinamento

Arquiteturas de SNDs	Operações Aritméticas	EMQ	Faixa de EMQ	Varição EMQ	Acertos (%)
NEFCLASS	<i>Max/Min</i>	1,22	1,30 à 1,22	0,002	82,78
SisND/Eps - HIROTA e PEDRYCZ	Soma/Produto Algébrico	0,51	0,87 à 0,28	0,057	51,37
	Soma/Produto de Hamacher	0,49	0,49 à 0,35	0,00001	0,0
	Soma/Produto de Einstein	0,23	0,29 à 0,18	0,002	87,70
	<i>Max/Min</i>	0,24	0,38 à 0,24	0,01	100

A partir do Quadro 6.1.1 observou-se que a arquitetura NEFCLASS, utilizando as operações aritméticas difusas padrões *max/min*, obteve o maior EMQ em relação às operações aritméticas difusas implementadas a partir da arquitetura do SisND/EPs.

Apesar das operações aritméticas difusas padrões *max/min*, implementadas com o neurônio difuso proposto por Hirota-Pedrycz apresentarem o maior percentual de acertos na fase de treinamento, essas operações aritméticas difusas foram excluídas do processo de avaliação por apresentarem problemas no critério de convergência: saída da RNA com resultados iguais e instabilidade, ou seja, a RNA se manteve instável, mesmo os pesos sendo alterados, a RNA não classificou corretamente os padrões.

Desta forma, as operações aritméticas difusas Soma/Produto de Einstein OU(E) e as operações aritméticas difusas padrões *max/min* obtiveram os maiores percentuais de acertos, respectivamente; os menores percentuais de acertos ficaram com as operações aritméticas difusas Soma/Produto Algébrico OU(E) e Soma/Produto de Hamacher OU(E), respectivamente.

6.2.3 Kwan-Cai

O neurônio difuso proposto por Kwan-Cai (1994) difere do neurônio difuso proposto Hirota-Pedrycz (1992, 1994) na operação de confluência:

(a) inicialmente ocorre um produtório entre os pesos sinápticos e os valores de entrada para depois serem aplicadas as respectivas operações aritméticas difusas *t-normas* ou *t-conormas*.

Após a implementação das operações aritméticas difusas Soma/Produto Algébrico OU(E), foram observados os seguintes resultados de EMQ por época (Figura 6.1.7).

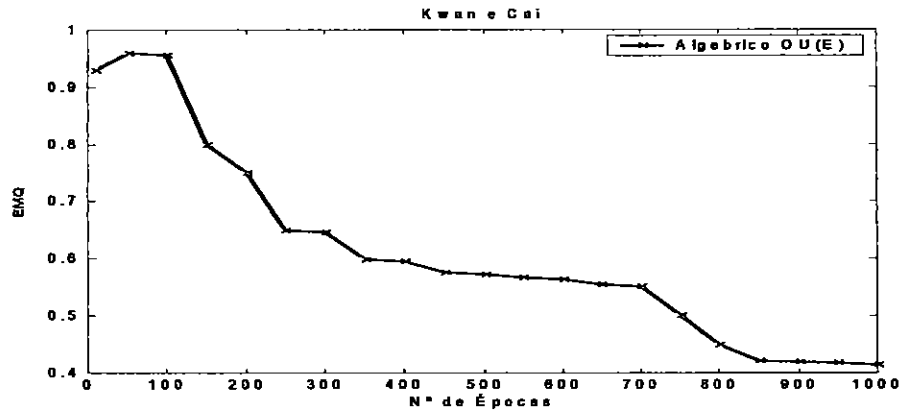


Figura 6.1.7 - EMQ - Soma/Produto Algébrico para o neurônio difuso Kwan- Cai

De acordo com a Figura 6.1.7, o EMQ na fase de treinamento permaneceu entre os valores $[0,92; 0,41]$, com uma média de EMQ de 0,55 e apresentando uma Variação da Taxa Absoluta do EMQ de 0,05. O percentual de acertos na fase de treinamento foi de 51,64%. Observou-se que o melhor resultado da aprendizagem ocorreu até a época 850, após o que a aprendizagem permaneceu estável até o final da fase de treinamento.

As operações aritméticas difusas Soma/Produto de Hamacher OU(E) apresentaram os resultados de EMQ por época mostrados na Figura 6.1.8.

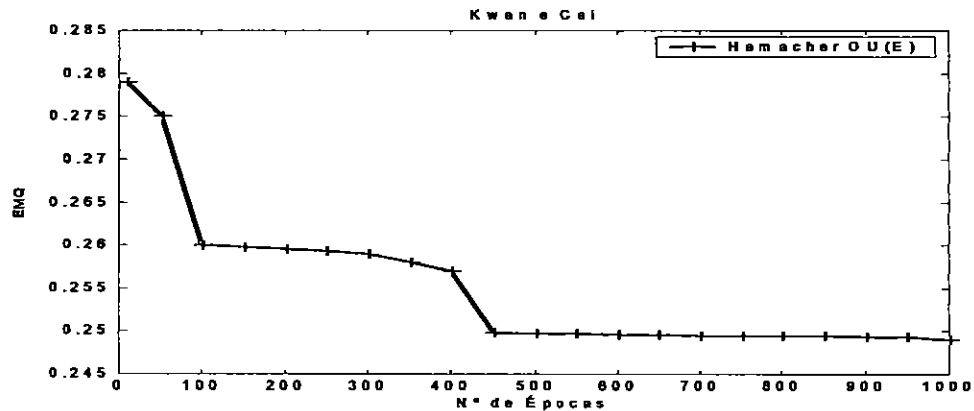


Figura 6.1.8 - RMS - Soma/Produto de Hamacher para o neurônio difuso Kwan-Cai

De acordo com a Figura 6.1.8, observou-se que o EMQ iniciou-se em 0,27, diminuindo até 0,25; após, permaneceu estável até o final do treinamento, apresentando altos percentuais de acertos na fase de treinamento, porém com problemas no critério de convergência. Observou-se que, próximo à época 450, ocorreu uma paralisia na modificação dos pesos sinápticos e a aprendizagem permaneceu estável até o final do treinamento.

As operações aritméticas difusas Soma/Produto de Einstein OU(E) apresentaram os resultados de EMQ por época mostrados na Figura 6.1.9.

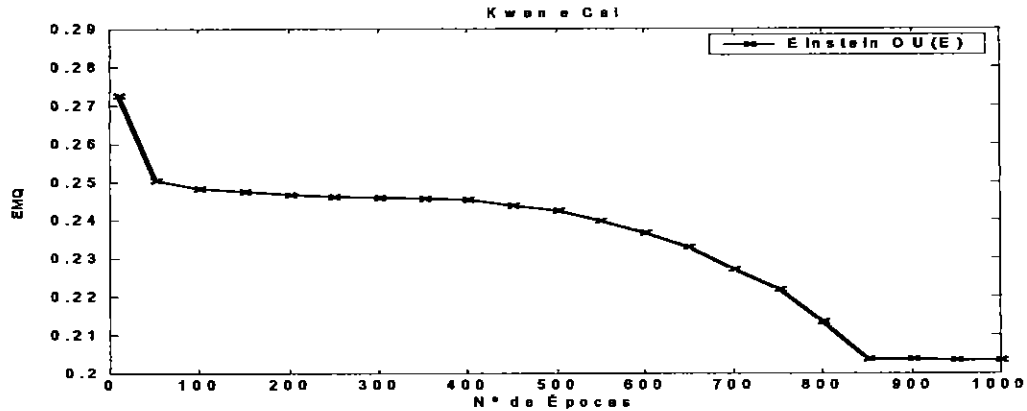


Figura 6.1.9 - EMQ - Soma/Produto de Einstein para o neurônio difuso Kwan-Cai

De acordo com a Figura 6.1.9, o EMQ permaneceu entre os valores $[0,27; 0,20]$, com uma média de EMQ de 0,21 e uma Variação da Taxa Absoluta do EMQ em 0,001. O percentual de acertos do treinamento foi de 93,44%. Observou-se que o EMQ minimizou gradualmente próximo à época 850; após, a aprendizagem permaneceu estável até o final do treinamento.

As operações aritméticas difusas padrões *max/min* apresentaram os resultados de EMQ por época mostrados na Figura 6.1.10.

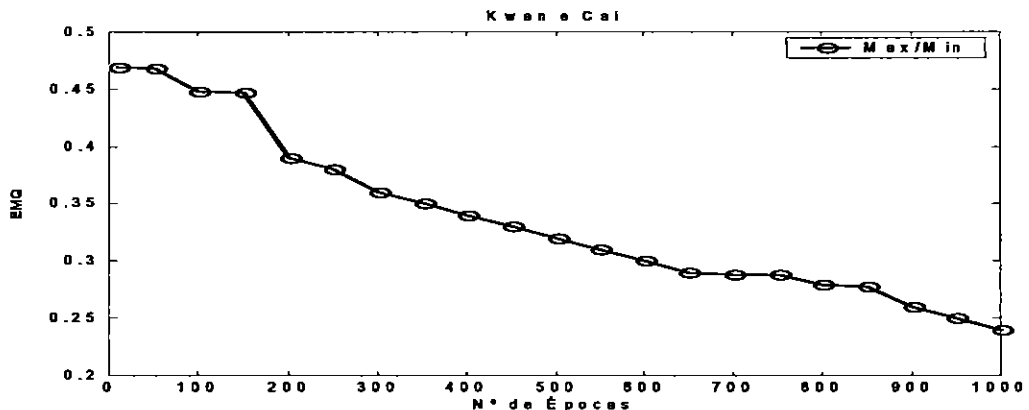


Figura 6.1.10 - EMQ – *Max/Min* para o neurônio difuso Kwan-Cai

De acordo com a Figura 6.1.10, o EMQ permaneceu entre os valores $[0,46; 0,24]$, com uma média de EMQ de 0,24 e uma Variação da Taxa Absoluta do EMQ em 0,01. O percentual de acertos no treinamento foi de 57,37%, porém a saída da RNA apresentou resultados iguais.

Observou-se que a aprendizagem ocorreu de forma gradual, com a minimização do EMQ até o final do treinamento.

Os resultados obtidos na aprendizagem por meio do neurônio difuso proposto por Kwan-Cai são mostrados na Figura 6.1.11.

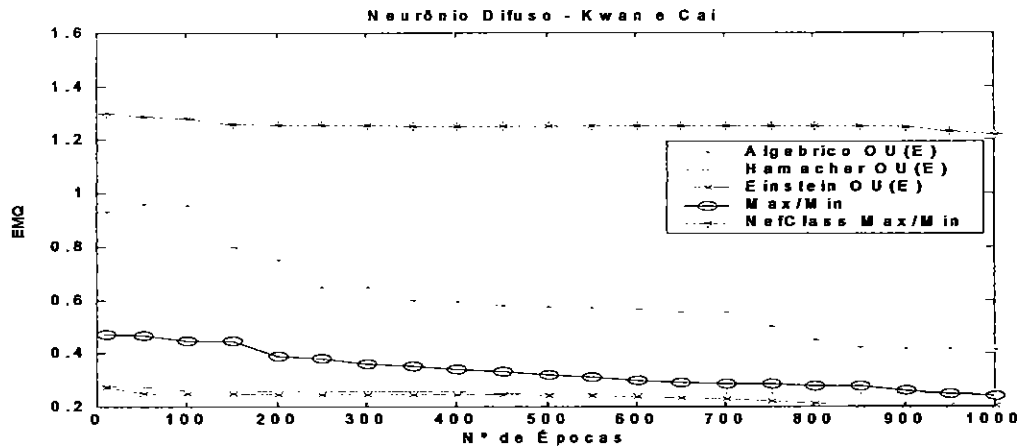


Figura 6.1.11 - EMQ da arquitetura NEFCLASS e do neurônio difuso Kwan-Cai

Observa-se a partir da Figura 6.1.11 que as operações aritméticas difusas padrões *max/min* implementadas com a arquitetura NEFCLASS continuaram com o maior EMQ por época em relação às demais operações aritméticas difusas; as operações aritméticas difusas Soma/Produto de Einstein OU(E) permaneceram com o menor EMQ por época. O maior percentual de acertos foi obtido a partir das operações aritméticas difusas Soma/Produto de Einstein OU(E), com 93,44%; o menor percentual de acertos (51,64%) ficou com as operações aritméticas difusas Soma/Produto Algébrico OU(E). A aprendizagem utilizando as operações aritméticas difusas Soma/Produto Algébrico OU(E) e Soma/Produto de Einstein OU(E) convergiu melhor em torno de 850 épocas de treinamento; após, permaneceu constante, ao passo que a aprendizagem obtida com a arquitetura NEFCLASS por meio das operações aritméticas difusas padrões *max/min* permaneceu constante a partir da época 350 até a época 850; após, o sistema voltou a convergir com a minimização do EMQ até o final do treinamento.

Após a avaliação no processo de aprendizagem, foram excluídas da análise estatística as operações aritméticas difusas Soma/Produto de Hamacher OU(E) (Figura 6.1.8), por apresentarem problemas na fase de treinamento, entre eles: EMQ estável, ocorrendo paralisia na modificação dos pesos e problemas no critério de convergência. Foram excluídas também da análise estatística as operações aritméticas difusas padrões *max/min* (Figura 6.1.10), ambas implementadas com o neurônio difuso proposto por Kwan-Cai, por apresentarem: saída da RNA com resultados iguais e

instabilidade, ou seja, a RNA se manteve instável, mesmo os pesos sendo alterados, a RNA não classificou corretamente os padrões.

Algumas razões que levaram a RNA a apresentar estes resultados, incluem:

(a) após a fase de treinamento com as operações aritméticas difusas padrões *max/min*, especificamente em relação à implementação da função sinal do gradiente (E^p), correspondente ao erro no padrão p , verificou-se que o valor da variável β precisa ser "negativo", caso contrário, a saída da RNA fornece resultados iguais;

(b) analisando os resultados das operações aritméticas difusas Soma/Produto de Hamacher OU(E), o problema da paralisia na modificação dos pesos ocorreu em razão das unidades de saída da RNA conterem valores próximos de "0" ou "1".

O Quadro 6.1.2 descreve os resultados das implementações na fase de treinamento obtidos com as arquiteturas NEFCLASS utilizando as operações aritméticas difusas padrões *max/min* e com o SisND/EPs por meio das operações aritméticas difusas Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein OU(E) e as operações aritméticas difusas padrões *max/min*. Os percentuais de acertos descritos referem-se aos acertos obtidos na fase de treinamento.

Quadro 6.1.2 - Desempenho das arquiteturas NEFCLASS e Kwan-Cai: fase de treinamento

Arquiteturas de SNDs	Operações Aritméticas	EMQ	Faixa de EMQ	Varição EMQ	Acertos (%)
NEFCLASS	<i>Max/Min</i>	1,22	1,30 à 1,22	0,002	82,78
SisND/EPs- KWANe CAI	Soma/Produto Algébrico	0,55	0,92 à 0,41	0,05	51,64
	Soma/Produto de Hamacher	0,43	0,27 à 0,24	0,0002	100,00
	Soma/Produto de Einstein	0,21	0,27 à 0,20	0,001	98,44
	<i>Max/Min</i>	0,24	0,46 à 0,24	0,01	100,00

A partir do Quadro 6.1.2 observa-se que a arquitetura NEFCLASS, por meio das operações aritméticas difusas padrões *max/min*, apresentou o maior EMQ por época em relação ao SisND/EPs; o menor EMQ foi obtido utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E). A menor faixa de EMQ foi obtida com as operações aritméticas difusas Soma/Produto de Einstein OU(E). Apesar das operações aritméticas difusas Soma/Produto de Hamacher OU(E),

juntamente com as operações aritméticas difusas padrões *max/min*, implementadas com o neurônio difuso Kwan-Cai, apresentarem o maior percentual de acertos na fase de treinamento, essas operações aritméticas foram excluídas do processo de avaliação por apresentarem problemas no critério de convergência.

Desta forma, os maiores percentuais de acertos foram obtidos por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E) e das operações aritméticas difusas padrões *max/min* implementadas com a arquitetura NEFCLASS, respectivamente. As operações aritméticas difusas Soma/Produto Algébrico OU(E) apresentaram o menor percentual de acertos (51,64%).

Após a exclusão das operações aritméticas difusas que não apresentaram resultados satisfatórios, inicia-se o processo de avaliação da aprendizagem com as arquiteturas NEFCLASS utilizando as operações aritméticas difusas padrões *max/min* e do SisND/EPs por meio das operações aritméticas difusas Soma/Produto Algébrico OU(E) e Soma/Produto de Einstein OU(E).

6.2.4 Desempenho das Arquiteturas NEFCLASS e do SisND/EPs

O Quadro 6.1.3 mostra os resultados finais obtidos com a comparação de desempenho das arquiteturas NEFCLASS e do SisND/EPs, os quais farão parte do processo de avaliação da aprendizagem, na fase de treinamento.

Quadro 6.1.3 - Comparação de desempenho das arquiteturas NEFCLASS e do SisND/EPs

Arquiteturas de SNDs		Operações Aritméticas	EMQ	Faixa de EMQ	Varição EMQ	Acertos (%)
NEFCLASS		Max/Min	1,22	1,30 à 1,22	0,002	82,78
S i s N D E P s	HIROTA e PEDRYCZ	Soma/Produto Algébrico	0,51	0,87 à 0,28	0,057	51,37
		Soma/Produto de Einstein	0,23	0,29 à 0,18	0,002	87,70
	KWAN e CAI	Soma/Produto Algébrico	0,55	0,92 à 0,41	0,05	51,64
		Soma/Produto de Einstein	0,21	0,27 à 0,20	0,001	93,44

A partir do Quadro 6.1.3 observam-se os seguintes resultados: o maior EMQ foi obtido com a arquitetura NEFCLASS; o menor EMQ ficou com as operações aritméticas difusas Soma/Produto de Einstein OU(E). O neurônio difuso proposto por Kwan-Cai obteve o melhor percentual de acertos (93,44%) utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E). Em segundo

lugar, com 87,70% de acertos, encontra-se o modelo implementado com o neurônio difuso proposto por Hirota-Pedrycz utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E). A arquitetura NEFCLASS ficou em terceiro lugar, com 82,78% de acertos com as operações aritméticas difusas padrões *max/min*. O menor percentual de acertos foi obtido com o neurônio difuso proposto por Kwan-Cai por meio das operações aritméticas difusas Soma/Produto Algébrico OU(E) (51,64%).

A seguir, iniciou-se a fase de avaliação de desempenho na qual a análise estatística foi aplicada aos melhores modelos de aprendizagem obtidos na fase de treinamento: (1) NEFCLASS utilizando as operações aritméticas difusas padrões *max/min*; (2) neurônios difusos propostos por Hirota-Pedrycz e Kwan-Cai por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E) e Soma/Produto Algébrico OU(E).

6.3 Avaliação de Desempenho das Arquiteturas NEFCLASS e do SisND/EPs

Como descrito no capítulo 5, o teste de hipóteses para proporções (%) verifica a consistência de uma hipótese estatística comparando o percentual de acertos da arquitetura do SisND/EPs com os resultados obtidos a partir da arquitetura NEFCLASS, adotando-se um nível de significância (α) de 0,05.

6.3.1 Teste de Hipótese das Arquiteturas NEFCLASS e do SisND/EPs

Os valores dos percentuais de acertos são apresentados no Quadro 6.1.4, considerando os 122 casos da fase de teste (n).

Quadro 6.1.4 - Teste de hipótese das arquiteturas NEFCLASS e do SisND/EPs

Arquiteturas de SNDs		Operações Aritméticas	n	Acertos (%)	p_valor
NEFCLASS		<i>Max/Min</i>	122	83,60	
S i s N D / E P s	HIROTA e PEDRYCZ	Soma/Produto de Einstein	122	86,06	0,2961
		Soma/Produto Algébrico	122	41,80	0,0001
	KWAN e CAI	Soma/Produto de Einstein	122	90,16	0,0646
		Soma/Produto Algébrico	122	43,44	0,0001

De acordo com o Quadro 6.1.4, ao comparar as arquiteturas do SisND/EPs com o NEFCLASS observou-se que as operações aritméticas difusas Soma/Produto Algébrico OU(E) utilizando os neurônios difusos Hirota-Pedrycz e Kwan-Cai apresentaram resultados piores do que os obtidos na arquitetura NEFCLASS. Os resultados obtidos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) com o neurônio difuso Hirota-Pedrycz foram semelhantes à arquitetura NEFCLASS ($p_valor > \alpha$). As operações aritméticas difusas Soma/Produto de Einstein OU(E) por meio do neurônio difuso Kwan-Cai apresentam a tendência de serem estatisticamente melhores do que o NEFCLASS ($p_valor \cong \alpha$).

A próxima etapa consiste na validação interna da base de conhecimentos utilizando as medidas de sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo.

6.4 Validação Interna

Realizou-se a validação interna da base de conhecimentos com as medidas estatísticas descritas no capítulo 5 e aplicadas aos neurônios difusos propostos por Kwan-Cai, Hirota-Pedrycz e a arquitetura NEFCLASS, respectivamente. Os valores que representaram o diagnóstico do sistema (*positivo/negativo*) foram obtidos a partir da fase de teste do SisND/EPs; os valores que representam o diagnóstico do especialista foram obtidos a partir do especialista de domínio.

6.4.1 Neurônio Difuso Kwan-Cai

O Quadro 6.1.5 mostra os resultados obtidos após a validação interna da base de conhecimentos utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E) implementadas com o neurônio difuso proposto por Kwan-Cai.

Quadro 6.1.5 - Validação interna - neurônio difuso Kwan-Cai

Diagnóstico do Sistema	Diagnóstico do Especialista		TOTAL
	Evento Epiléptico (positivo)	Evento não Epiléptico (negativo)	
Evento Epiléptico (positivo)	44	3	47 VP_p
Evento não Epiléptico (negativo)	9	66	75 VP_n
TOTAL	53 SE_n	69 ES_p	122

De acordo com o Quadro 6.1.5, a sensibilidade do sistema de detectar que um paciente possui um evento epiléptico foi de 83,01%, ou seja, 44 diagnósticos apresentando eventos epilépticos tenderam a um ponto comum, ou seja, convergiram e 9 diagnósticos de eventos epilépticos não convergiram; a especificidade do sistema de detectar que um paciente possui um evento não epiléptico foi de 95,65%, isto é, 66 diagnósticos apresentando eventos não epilépticos convergiram e 3 diagnósticos de eventos não epilépticos não convergiram.

Em relação ao caso ser um evento epiléptico quando o resultado da classificação disponibilizado pelo sistema for positivo, os resultados mostraram um valor preditivo positivo de 93,61%; em relação ao caso ser um evento não epiléptico quando o resultado da classificação disponibilizado pelo sistema for negativo, o valor preditivo negativo encontrado foi de 88%.

6.4.2 Neurônio Difuso Hirota-Pedrycz

O Quadro 6.1.6 mostra as operações aritméticas difusas Soma/Produto de Einstein OU(E) obtidas com o SisND/EPs a partir do neurônio difuso proposto por Hirota-Pedrycz.

Quadro 6.1.6 - Validação interna - neurônio diuso Hirota-Pedrycz

Diagnóstico do Sistema	Diagnóstico do Especialista		TOTAL
	Evento Epiléptico (positivo)	Evento não Epiléptico (negativo)	
Evento Epiléptico (positivo)	42	6	48
Evento não Epiléptico (negativo)	11	63	74
TOTAL	53	69	122

A partir do Quadro 6.1.6 observou-se que a sensibilidade e a especificidade obtiveram valores menores em relação ao modelo proposto por Kwan-Cai. A sensibilidade do sistema de detectar que um paciente possui um evento epiléptico ficou em 79,24%, ou seja, 42 diagnósticos apresentando eventos epilépticos convergiram e 11 diagnósticos de eventos epilépticos não convergiram; a especificidade do sistema de detectar que um paciente possui um evento não epiléptico diminuiu para

91,30%, ou seja, 63 diagnósticos apresentando eventos não epiléticos convergiram e 6 diagnósticos de eventos não epiléticos não convergiram.

Em relação ao caso implementado ser um evento epilético quando o resultado da classificação disponibilizado pelo sistema for positivo, os resultados mostraram um valor preditivo positivo de 87,50%; em relação ao caso ser um evento não epilético quando o resultado da classificação disponibilizado pelo sistema for negativo, o valor preditivo negativo obtido foi de 85,13%.

6.4.3 Arquitetura NEFCLASS - Operações Aritméticas Difusas Padrões *Max/Min*

O Quadro 6.1.7 ilustra as operações aritméticas difusas padrões *max/min* implementadas a partir da arquitetura *NEFCLASS*.

Quadro 6.1.7 - Validação interna - operações aritméticas difusas padrões *Max/Min*

Diagnóstico do Sistema	Diagnóstico do Especialista		TOTAL
	Evento Epilético (positivo)	Evento não Epilético (negativo)	
Evento Epilético (positivo)	45	12	57 VP_p
Evento não Epilético (negativo)	8	57	65 VP_n
TOTAL	53 SE_n	69 ES_p	122

Observou-se, a partir do Quadro 6.1.7, que apenas a sensibilidade do sistema obteve percentual de acerto maior em relação aos demais modelos:

- a sensibilidade do sistema de detectar que um paciente possui um evento epilético obteve 84,90%, ou seja, 45 diagnósticos apresentando eventos epiléticos convergiram e 8 diagnósticos de eventos epiléticos não convergiram;

- a especificidade do sistema de detectar que um paciente possui um evento não epilético ficou em 82,60%, isto é, 57 diagnósticos apresentando eventos não epiléticos convergiram e 12 diagnósticos de eventos não epiléticos não convergiram.

Em relação ao caso ser um evento epilético quando o resultado da classificação disponibilizado pelo sistema for positivo, os resultados mostraram um valor preditivo positivo de 78,94%; em relação

ao caso ser um evento não epiléptico quando o resultado da classificação disponibilizado pelo sistema for negativo, o valor preditivo negativo foi de 87,69%.

6.4.4 Sensibilidade e Valor Preditivo Positivo

Após a aplicação do teste de hipóteses para proporções ³², que verifica a consistência de uma hipótese estatística, comparando o percentual de acertos da arquitetura do SisND/EPs, através das operações aritméticas difusas Soma/Produto de Einstein OU(E), e utilizando os neurônios difusos propostos por Hirota-Pedrycz e Kwan-Cai com os resultados obtidos a partir da arquitetura NEFCLASS, a qual utiliza as operações aritméticas difusas padrões *max/min*, os valores obtidos com os procedimentos de teste a partir dos percentuais de acertos das variáveis sensibilidade e valor preditivo positivo são apresentados no Quadro 6.1.8.

Foram considerados os 122 casos da base de teste.

Quadro 6.1.8 - Validação interna - sensibilidade e valor preditivo positivo

Arquiteturas de SNDs		Operações Aritméticas	n	SEn (%)	p_valor	VPp (%)	p_valor
NEFCLASS		<i>Max/Min</i>	122	84,90	-	78,94	-
S i s N D / E P s	HIROTA e PEDRYCZ	Soma/Produto de Einstein	122	79,24	0,1246	87,50	0,0368
	KWAN e CAI	Soma/Produto de Einstein	122	83,01	0,3438	93,61	0,00043

Como se pode observar, a partir do Quadro 6.1.8, a melhor sensibilidade do sistema foi obtida com a arquitetura NEFCLASS utilizando as operações aritméticas difusas padrões *max/min*, ou seja, a capacidade do sistema de detectar casos positivos da hipótese diagnóstica foi de 84,90%. Em relação ao valor preditivo positivo, os melhores valores contemplaram as operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz, respectivamente.

Ao comparar as arquiteturas do SisND/EPs com o NEFCLASS, aplicando o teste de hipóteses para proporções, observou-se que, utilizando o neurônio difuso proposto por Hirota-Pedrycz por meio das operações aritméticas Soma/Produto de Einstein OU(E), os resultados de sensibilidade

³² O teste de hipóteses para proporções é um método que utiliza os dados da amostra para decidir se a hipótese nula deve ser desprezada ou não (Devore, 2006).

são piores do que os obtidos na arquitetura NEFCLASS ($p_valor > \alpha$); note-se que no neurônio difuso proposto por Kwan-Cai a sensibilidade foi semelhante à arquitetura NEFCLASS por meio das operações aritméticas difusas padrões max/min ($p_valor > \alpha$). Os resultados de valor preditivo positivo utilizando os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz apresentaram-se estatisticamente melhores do que a arquitetura NEFCLASS, que utilizou as operações aritméticas difusas padrões max/min ($p_valor < \alpha$), respectivamente.

6.4.5 Especificidade e Valor Preditivo Negativo

Em relação à especificidade, isto é, à capacidade do sistema de detectar casos negativos da hipótese diagnóstica, o modelo de aprendizagem baseado no neurônio difuso proposto por Kwan-Cai obteve o melhor resultado (95,65%) em relação aos 91,30% obtidos com o neurônio difuso proposto por Hirota-Pedrycz, ambos com as operações aritméticas difusas Soma/Produto de Einstein OU(E). As operações aritméticas difusas padrões max/min implementadas com a arquitetura NEFCLASS obtiveram 82,60% de especificidade.

Os valores dos percentuais de acertos obtidos após a validação interna da base de conhecimentos em relação à especificidade e ao valor positivo negativo são mostrados no Quadro 6.1.9.

Quadro 6.1.9 - Validação interna - especificidade e valor preditivo negativo

Arquiteturas de SNDs		Operações Aritméticas	n	ESp (%)	p_valor	VPn (%)	p_valor
NEFCLASS		<i>Max/Min</i>	122	82,60	-	87,69	-
S i s N D / E P s	HIROTA e PEDRYCZ	Soma/Produto de Einstein	122	91,30	0,0218	85,13	0,2798
	KWAN e CAI	Soma/Produto de Einstein	122	95,65	0,00053	88,00	0,4705

Finalmente, quanto ao valor preditivo negativo, ou seja, os pacientes não pertencerem à hipótese diagnóstica quando o resultado for negativo, os três modelos apresentaram resultados semelhantes ($p_valor > \alpha$).

De acordo com o Quadro 6.1.9, após a aplicação do teste de hipóteses para proporções, ao comparar as arquiteturas do SisND/EPs com o NEFCLASS, observou-se que os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz são estatisticamente melhores em especificidade do que o

NEFCLASS ($p_valor < \alpha$), respectivamente; note-se que o valor preditivo negativo obtido com os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz apresentou resultados semelhantes ao NEFCLASS ($p_valor > \alpha$).

6.5 Comportamento da Aprendizagem

A arquitetura do SisND/EPs apresentou diferenças no comportamento da aprendizagem quando aplicados diferentes ajustes nas variáveis que compõem o sistema.

Na seqüência, os resultados obtidos com o comportamento da aprendizagem das arquiteturas NEFCLASS, SisND/EPs, sensibilidade, especificidade, valor preditivo positivo e negativo.

6.5.1 NEFCLASS e do SisND/EPs

O Quadro 6.1.10 mostra os resultados obtidos após a alteração das variáveis: (a) pesos sinápticos (neurônios difusos E e OU , respectivamente); (b) valor do termo $bias$; (c) termo "momento"; (d) EMQ; (e) percentual de acertos. O valor da taxa de aprendizagem permaneceu em 0,01; pesos aleatórios no intervalo $[-0,5$ até $0,5]$; pesos fixos em 0,01 e valor do $bias$ (-1). Os percentuais de acertos descritos referem-se aos acertos obtidos na fase de teste.

Foram executadas três fases de treinamento a partir dos modelos implementados, sendo que os valores da primeira fase de treinamento correspondem aos utilizados na aprendizagem dos modelos (seção 6.2) no início deste capítulo. Os valores utilizados na arquitetura NEFCLASS representam apenas os valores da primeira fase de treinamento pois são desconhecidos da literatura os valores de ajustes do valor do $bias$ e do termo "momento".

A arquitetura NEFCLASS, apesar de ser um software disponível na Web, apresenta limitações, ou seja, o seu código de máquina é inacessível ao usuário, não permitindo alterações nos parâmetros do programa.

No Quadro 6.1.10 são descritos os resultados apresentados pelas melhores operações aritméticas difusas:

- (a) operações aritméticas difusas padrões max/min , implementadas com a arquitetura NEFCLASS;
- (b) operações aritméticas difusas Soma/Produto Algébrico OU(E);
- (c) operações aritméticas difusas Soma/Produto de Einstein OU(E), estas duas últimas, implementadas com os neurônios difusos propostos por Hirota-Pedrycz e Kwan-Cai.

O aprendizado foi definido em relação ao percentual de acertos obtidos na fase de teste.

Quadro 6.1.10 - Comportamento da aprendizagem - NEFCLASS e do SisND/EPs

Arquiteturas	Operações Aritméticas	Pesos		Bias	Momento		EMQ			Acertos (%)	Aprendizado	
		Aleat.	Fixos	Negativo	0,85	0,95	Faixa	Varição	Média			
NEFCLASS	<i>Max/Min</i>	1ª	-	1	-	-	1,30 – 1,22	0,002	1,19	83,60	Adequado	
HIROTA e PEDRYCZ	Soma/Produto Algébrico OU(E)	1ª	X	-	Positivo	X	-	0,87 – 0,28	0,57	0,51	41,80	Inadequado
		2ª	X	-	X	X	-	0,82 – 0,41	0,021	0,53	45,08	Inadequado
		3ª	-	X	X	X	-	0,82 – 0,41	0,021	0,53	45,08	Inadequado
	Soma/Produto de Einstein OU(E)	1ª	X	-	Positivo	X	-	0,29 – 0,18	0,002	0,23	86,06	Adequado
		2ª	X	-	X	X	-	0,28 – 0,18	0,005	0,23	80,33	Adequado
		3ª	-	X	X	-	X	0,27 – 0,25	0,001	0,25	56,55	Inadequado
KWAN e CAI	Soma/Produto Algébrico OU(E)	1ª	X	-	Positivo	X	-	0,92 – 0,41	0,052	0,51	43,44	Inadequado
		2ª	X	-	X	-	X	0,77 – 0,35	0,018	0,46	45,08	Inadequado
		3ª	X	-	Positivo	X	-	0,93 – 0,41	0,027	0,52	42,62	Inadequado
	Soma/Produto de Einstein OU(E)	1ª	X	-	Positivo	X	-	0,27 – 0,20	0,001	0,21	90,16	Adequado
		2ª	X	-	X	X	-	0,29 – 0,21	0,004	0,24	56,56	Inadequado
		3ª	X	-	X	-	X	0,37 – 0,18	0,010	0,23	72,13	Inadequado

Observa-se no Quadro 6.1.10 que o maior EMQ foi obtido com a arquitetura NEFCLASS utilizando as operações aritméticas difusas padrões *max/min*, apresentando um percentual de acertos na fase de teste de 83,60%, o que caracterizou um aprendizado adequado. As operações aritméticas difusas Soma/Produto Algébrico OU(E) propostas pelos neurônios difusos Hirota-Pedrycz e Kwan-Cai continuaram com o menor percentual de acertos na fase de testes em relação às demais operações aritméticas difusas, caracterizando um aprendizado inadequado. Os melhores percentuais de acertos na fase de teste e também os menores índices de EMQ foram obtidos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz, caracterizando-os como aprendizados adequados. A menor variação da Taxa Absoluta de EMQ médio foi obtido por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando o neurônio difuso proposto por Kwan-Cai.

A partir do comportamento da aprendizagem dos modelos, pode-se observar que, após a alteração dos valores referentes à segunda e terceira fase, os percentuais de acertos na fase de teste permaneceram com baixos índices, caracterizando-os como aprendizados inadequados, com exceção dos resultados obtidos na segunda fase de treinamento a partir do neurônio difuso proposto por

Hirota-Pedrycz, utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E), o que caracterizou-o como aprendizado adequado.

6.5.2 Sensibilidade, Especificidade, Valor Preditivo Positivo e Negativo

Em relação às medidas de sensibilidade, especificidade, valor preditivo positivo e valor preditivo negativo, o comportamento da aprendizagem é mostrado no Quadro 6.1.11.

Quadro 6.1.11 - Aprendizagem - sensibilidade, especificidade, valor preditivo positivo e negativo

Arquiteturas	Operações Aritméticas	SEn (%)	Aprendizado	ESp (%)	Aprendizado	VPp (%)	Aprendizado	VPn (%)	Aprendizado
NEFCLASS	<i>Maz/Min</i>	84,90	Adequado	82,60	Adequado	78,94	Inadequado	87,69	Adequado
HIROTA e PEDRYCZ	Soma/Produto de Einstein OU(E)	79,24	Inadequado	91,30	Adequado	87,50	Adequado	85,13	Adequado
KWAN e CAI	Soma/Produto de Einstein OU(E)	83,01	Adequado	95,65	Adequado	93,61	Adequado	88,00	Adequado

Observa-se no Quadro 6.1.11 que os melhores resultados relacionados com a sensibilidade foram obtidos utilizando a arquitetura NEFCLASS e com o neurônio difuso proposto por Kwan-Cai, caracterizando-os como aprendizados adequados. Note-se que na arquitetura NEFCLASS o valor preditivo positivo foi considerado inadequado (78,94%). Os melhores resultados de especificidade foram obtidos por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando os neurônios difusos propostos por Kwan-Cai, Hirota-Pedrycz e com a arquitetura NEFCLASS, caracterizando-os como aprendizados adequados, respectivamente. Os melhores resultados de valor preditivo positivo foram obtidos a partir dos neurônios difusos Kwan-Cai e Hirota-Pedrycz, caracterizando-os como aprendizados adequados, respectivamente. Nos resultados de valor preditivo negativo todos os modelos implementados foram considerados adequados.

Observando a avaliação geral das arquiteturas implementadas, conclui-se que a implementação por meio dos neurônios difusos, apresenta influência positiva no desempenho do SisND/EPs, principalmente em relação ao percentual de acertos e faixa de EMQ³³, como mostra o Quadro 6.1.10.

33

As arquiteturas implementadas, utilizando neurônios difusos, iniciaram com o menor EMQ.

6.6 Considerações Finais

As arquiteturas implementadas para explorar diferentes operações aritméticas difusas foram baseadas no hibridismo integrando RNAs e a LD.

Com base nos resultados apresentados observa-se que as operações aritméticas difusas Soma/Produto de Einstein OU(E) implementadas com os neurônios difusos propostos por Kwan-Cai, Hirota-Pedrycz, juntamente com a arquitetura NEFCLASS, por meio das operações aritméticas difusas padrões *max/min*, são modelos de aprendizagem em SNDs e apresentaram os melhores percentuais de acertos na fase de teste: 90,16%; 86,06% e 83,60%, respectivamente.

A avaliação de desempenho mostra que as operações aritméticas difusas Soma/Produto de Einstein OU(E) implementadas por meio do neurônio difuso Hirota-Pedrycz apresentaram resultados semelhantes à arquitetura NEFCLASS; as operações aritméticas difusas Soma/Produto de Einstein OU(E) implementadas com o neurônio difuso Kwan-Cai apresentam a tendência de serem estatisticamente melhores do que as da arquitetura NEFCLASS, com a probabilidade de significância de 0,06.

Na validação interna da base de conhecimentos observa-se que o neurônio difuso proposto por Kwan-Cai utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E) e a arquitetura NEFCLASS obtiveram os melhores resultados para classificar tanto os eventos epiléticos como os eventos não epiléticos, ao passo que o neurônio difuso proposto por Hirota-Pedrycz obteve melhor resultado na classificação dos eventos não epiléticos.

Foram confrontados com a literatura apenas os resultados relacionados com a arquitetura NEFCLASS, pois até o momento não foram encontradas referências na literatura consultada relacionadas com a exploração das diferentes operações aritméticas difusas propostas nesta tese. Os resultados encontrados utilizando as operações aritméticas difusas padrões *max/min*, por meio da arquitetura NEFCLASS, comprovam uma similaridade, mostrando percentuais de acertos compatíveis com a literatura. Conde, Ramos e Vasconcelos (2006) investigaram dois SNDs: (1) câncer de mama e (2) diabetes mellitus, os quais apresentaram um percentual de acertos de 98,28% e 76,72%, respectivamente. Serhatlioglu et al. (2004, 2003) investigaram diabetes mellitus em artérias oftálmicas (direita e esquerda) e nas artérias carótidas, obtendo percentuais de acertos de 85,0%, 87,5% e 85,0%, respectivamente. Nauck e Kruse (1997_a) exploraram um SND para diagnosticar câncer de mama, apresentando um percentual de acertos de 96,5%. Em relação aos percentuais de sensibilidade e especificidade relacionados com um SND para diagnosticar glaucomas visuais, foram

encontrados na literatura consultada percentuais de acertos de 96,0% e 93,4% (Garcia-Feijoo et al., 2002).

No próximo capítulo são apresentadas as conclusões e sugestões de trabalhos futuros.

Capítulo 7

Conclusões

Esta tese avaliou e validou modelos de aprendizagem em SNDs utilizando diferentes operações aritméticas difusas *t-normas* e *t-conormas*, juntamente com diferentes tipos de neurônios difusos.

As arquiteturas NEFCLASS e o SisND/EPs permitiram modelar a incerteza por imprecisão presente na entrada da RNA por meio da implementação das operações aritméticas difusas.

Os resultados obtidos demonstraram a capacidade das arquiteturas de trabalharem com diferentes configurações de neurônios difusos, fornecendo resultados diferenciados na classificação de padrões. O desempenho das arquiteturas foi avaliado por meio de uma base de casos envolvendo os eventos epiléticos e os eventos não epiléticos.

Observou-se, a partir da revisão da literatura, que a maioria dos SNDs pesquisados utiliza o método de inferência baseado nas operações aritméticas difusas padrões *max/min*. Dessa forma, os SNDs investigados nesta tese distinguem-se dos modelos existentes por explorarem diferentes operações aritméticas difusas, como Soma/Produto Algébrico OU(E), Soma/Produto de Hamacher OU(E), Soma/Produto de Einstein OU(E) e as operações aritméticas difusas padrões *max/min* na composição dos neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz.

A partir da exploração da arquitetura do SisND/EPs, implementada com os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E), juntamente com a arquitetura NEFCLASS, foi possível obter resultados mais eficientes que os gerados com as operações aritméticas difusas Soma/Produto Algébrico OU(E) e Soma/Produto de Hamacher OU(E), tanto na avaliação de desempenho como na validação interna da base de conhecimentos. A partir desta investigação, é provável que estes resultados estejam relacionados com as diferenças encontradas nos "pontos de instanciações" das operações aritméticas difusas implementadas. Contudo, para comprovação estatística, sugere-se que sejam investigadas outras operações aritméticas difusas, juntamente com diferentes "pontos de instanciações".

Observou-se que os melhores percentuais de acertos na fase de teste foram obtidos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando os neurônios difusos

propostos por Kwan-Cai, Hirota-Pedrycz e da arquitetura NEFCLASS com as operações aritméticas difusas padrões *max/min*, caracterizando-os como aprendizados adequados, respectivamente.

Na avaliação de desempenho comparando o SisND/EPs com a arquitetura NEFCLASS observou-se que as operações aritméticas difusas Soma/Produto Algébrico OU(E) por meio dos neurônios difusos Hirota-Pedrycz e Kwan-Cai apresentaram resultados inferiores aos obtidos com a arquitetura NEFCLASS. Os resultados obtidos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando o neurônio difuso proposto por Hirota-Pedrycz não são estatisticamente diferentes do NEFCLASS, apresentando resultados semelhantes, ao passo que os resultados obtidos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) utilizando o neurônio difuso proposto por Kwan-Cai apresentaram a tendência de serem estatisticamente melhores do que a arquitetura NEFCLASS, com uma probabilidade de significância de 0,06.

A partir dos resultados apresentados na validação interna da base de conhecimentos, os melhores resultados de sensibilidade e valor preditivo positivo foram obtidos com o neurônio difuso proposto por Kwan-Cai utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E). Os melhores resultados de especificidade e valor preditivo negativo foram obtidos utilizando os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz por meio das operações aritméticas difusas Soma/Produto de Einstein OU(E).

Constatou-se, portanto, que os modelos de aprendizagem utilizando as operações aritméticas difusas padrões *max/min*, implementadas com a arquitetura NEFCLASS, juntamente com as operações aritméticas difusas Soma/Produto de Einstein OU(E), implementadas com o neurônio difuso proposto por Kwan-Cai, são melhores no diagnóstico de eventos epiléticos; os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz são melhores na classificação de eventos não epiléticos.

Considerando a avaliação geral, os melhores percentuais de acertos foram obtidos com: (1) Kwan-Cai e (2) Hirota-Pedrycz, ambos com as operações aritméticas difusas Soma/Produto de Einstein OU(E) e (3) arquitetura NEFCLAS por meio das operações aritméticas difusas padrões *max/min*. Considerando, ainda, a análise de sensibilidade e especificidade, conclui-se que as operações aritméticas difusas Soma/Produto de Einstein OU(E) implementadas com o neurônio difuso Kwan-Cai obtiveram um diagnóstico mais preciso.

Dessa forma, sugere-se:

(1) que as arquiteturas NEFCLASS, utilizando as operações aritméticas difusas padrões *max/min* e o SisND/EPs, implementado com o neurônio difuso proposto por Kwan-Cai, sejam utilizadas em paralelo no diagnóstico complementar de eventos epiléticos;

(2) que a arquitetura do SisND/EPs, utilizando as operações aritméticas difusas Soma/Produto de Einstein OU(E), implementada com os neurônios difusos propostos por Kwan-Cai e Hirota-Pedrycz, seja utilizada em paralelo no diagnóstico complementar dos eventos não epiléticos.

Concluindo, os resultados desta pesquisa, após o processo de avaliação e validação, poderão ser aplicados a outros SNDs com o objetivo de desenvolver sistemas com melhores percentuais de acertos. Existem várias aplicações, as quais podem utilizar os modelos de aprendizagem explorados nesta tese, por exemplo, outros domínios de aplicações de diagnóstico médico e reconhecimento de padrões.

Como trabalhos futuros na área de aprendizagem de SNDs faz-se algumas sugestões:

(1) fazer a validação concorrente com as arquiteturas NEFCLASS e o SisND/EPs em diferentes contextos clínicos e em áreas distintas do conhecimento, utilizando tanto as operações aritméticas difusas como os neurônios difusos explorados nesta tese nas etapas de avaliação da aprendizagem, avaliação de desempenho e na validação interna da base de conhecimentos;

(2) avaliar o desempenho de diferentes operações aritméticas difusas, incluindo *Lukasiewicz*, *Kleene-Dines*, *Yager*, etc;

(3) aplicar diferentes modelos de neurônios difusos, incluindo *Yamakawa* e *Kartalopoulos*.

Referências

- Aaai, American Association for Artificial Intelligence. Disponível em: <<http://www.aaai.org>>. Acesso em: 20 jan. 2007.
- Akcayol, M. A.; Elmas, C. NEFCLASS-based neuro fuzzy controller for SRM drive. *Engineering Applications of Artificial Intelligence*, v. 18, n. 5, p. 595-602, 2005.
- Ackley, D. H.; Hinton, G. E.; Sejnowski, T. J. Introduction to A Learning Algorithm for Boltzmann Machines. In: *Neurocomputing - Foundations of Research*. Cambridge, MA: Mit Press, Anderson and Rosenfeld (eds), p. 363, 1988.
- Aguiar, L.; Melin, P.; Castillo, O. Intelligent control of a stepping motor drive using a hybrid neuro-fuzzy ANFIS Approach. *Applied Soft Computing*, v. 3, n. 3, p. 209-219, nov. 2003.
- Aisb. Society for Artificial Intelligence and Simulation of Behaviour. Disponível em: <<http://www.aisb.org.uk>>. Acesso em: 16 jan. 2007.
- Amano, A.; Aritsuka, T.; Hataoka, N.; Ichikawa, A. On the use of neural network and fuzzy logic in speech recognition. In *Proc. of 1989 Int'l Joint Conf. on Neural Networks*, San Diego, v. 1, p. 301-305, 1989.
- Appel, E. Uma Proposta de Modelagem de Conhecimentos sobre Glaucoma por um Sistema Híbrido Neuro-Fuzzy. 2004. Tese (Doutorado), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2004.
- Argoud, F. I. M. Contribuição à Automatização da Detecção e Análise de Eventos Epileptiformes em EEG. 2001. Tese (Doutorado em Engenharia Elétrica) - Curso de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2001.
- Aual. Association for Uncertainty in Artificial Intelligence. Disponível em: <<http://www.auai.org>>. Acesso em: 10 jan. 2007.
- Avanzini, G.; 2003. Of cabbages and kings: Do we really need a systematic classification of epilepsies?, *Epilepsia*, v. 44, p. 12-13.
- Barr, A.; Feigenbaum. *The Handbook of Artificial Intelligence*. California: Willian Kaufmann Inc., v. 1, 1981.
- Barreto, J. M. *Inteligência Artificial - No limiar do século XXI*. 3. Ed. Florianópolis: Impresso por Duplic - Prestação de Serviços, 2001.
- Barreto, J. M. *Conexionismo e a Resolução de Problemas*. 1996. Tese (Concurso para Professor Titular), Departamento de Informática e Estatística, Universidade Federal de Santa Catarina, Florianópolis, 1996.

- Barreto, J. M.; De Azevedo, F. M. Connectionist expert systems as medical decision aid. *AI in Medicine*, v. 5, n. 9, p. 1-9, 1983.
- Battistella, E. Extração de Regras de Redes Neurais Artificiais Aplicadas ao Problema da Previsão da Estrutura Secundária de Proteínas. 2004. Dissertação (Mestrado), Universidade do Vale do Rio dos Sinos, São Leopoldo, 2004.
- Berenji H.R.; Khedkar, P. Learning and Tuning Fuzzy Logic Controllers Through Reinforcements, *IEEE Trans. Neural Networks*, v. 3, n. 5, p. 724-740, 1992.
- Berger, A. T.; Blackstone, N.W. Of cabbages and kings: Perspectives on classification from the field of systematics. *Epilepsia*, v. 44, p. 8-12, 2003.
- Bittencourt, G. Inteligência Artificial: ferramentas e teorias. Florianópolis: Ed. da Universidade Federal de Santa Catarina, 1998.
- Braga, M. J. F.; Barreto, J. M.; Machado, M. A. S. Conceitos da Matemática Nebulosa na Análise de Risco. Artes & Rabisus Comunicação Empresarial Ltda., Rio de Janeiro/RJ, 1995.
- Braga, A. P.; De Carvalho, A. C. P. L. F.; Ludermir, T. B. Redes Neurais Artificiais: Teoria e Aplicações. Rio de Janeiro: Livros Técnicos e Científicos Editora S. A. 2000.
- Brasil, L. M.; De Azevedo, F. M. Técnica de Extração de Regras para Sistemas Especialistas Conexionistas. Rio de Janeiro: Revista Brasileira de Engenharia Biomédica, v. 19, p. 7-17, 2003.
- Brasil, L. M. Proposta de Arquitetura para Sistema Especialista Híbrido e a Correspondente Metodologia de Aquisição do Conhecimento. 1999. Tese (Doutorado em Engenharia Elétrica), Universidade Federal de Santa Catarina, Florianópolis, 1999.
- Brasil, L. M. Aquisição de Conhecimento Aplicada ao Diagnóstico de Epilepsia. 1994. Dissertação (Mestrado em Engenharia Elétrica), Universidade Federal de Santa Catarina, Florianópolis, 1994.
- Brasil, L. M.; De Azevedo, F. M.; Barreto, J. M. A Hybrid Expert Systems for the Diagnosis of Epileptic Crisis. *Artificial Intelligence in Medicine, Elsevier Science*, v. 21, p. 227-233, 2001.
- Brawn, T. H.; Kairiss, E. W.; Keenan, C. L. Hebbian synapses: Biophysical mechanisms and algorithms, *Annual Review of Neuroscience*, v. 13, p. 475-511, 1990.
- Caudill, M.; Butler, C. *Understanding Neural Networks*. v. 1. Basic Networks, USA: Massachusetts Institute of Technology, 1992_a.
- Caudill, M.; Butler, C. *Understanding Neural Networks*. v. 2. Advanced Networks, USA: Massachusetts Institute of Technology, 1992_b.
- Cechin, A. L. The Extraction of Fuzzy Rules from Neural Networks. 1998. Shaker Verlag. Aachen, Alemanha. Tese de Doutorado - Universidade de Tübingen, 1998.

- Cechin, A. L.; Osório, F. S. KD - O Conhecimento: Formas de Representação e Exploração. Cadernos de Informática, v. 1, n. 1, p. 7-13, 2001.
- Cecil M. Medicina Interna Básica. Ed. Guanabara: Rio de Janeiro, 1990.
- Conde, G. A.; Ramos, P. G.; Vasconcelos, G. C. Neuro-Fuzzy Networks for Pattern Classification and Rule Extraction. In: Simposio Brasileiro de Redes Neurais, 2000. Disponível em: <<http://csdl.computer.org/comp/proceedings/sbrn/2000/0856/00/08560289.pdf>>. Acesso em: 25 set. 2006.
- Conway, O. M.; Mcgonigal, P.; Russel, A., et al. Differences in psychogenic nonepileptic seizures. Disponível em: <<http://www.sciencedirect.com>>. Acesso em: 07 dez. 2006.
- Costa, Da J. C.; Palmimi, A.; Yacubian, El M. T.; Cavalheiro, E. A. Fundamentos Neurobiológicos das Epilepsias: aspectos clínicos e cirúrgicos, v. 1, São Paulo: Lemos Editorial, 1988.
- Costa, P. A. B. Um Enfoque segundo a Teoria de Conjuntos Difusos para a Meta-Análise. 1999. Florianópolis. Tese (Doutorado em Engenharia da Produção), Universidade Federal de Santa Catarina, SC, 1999.
- De Azevedo, F. M.; Barreto, J. M.; Lima, W. C. de. A neural network approach for fuzzy knowledge bases. In: Proceedings of the IX Congreso de Ingeniería Eléctrica, p. 9.1.1 - 9.1.6, Chile: Universidad de Tarapaca, 1988.
- De Azevedo, F. M. Contribution to the Study of Neural Networks in Dynamical Expert System. 1993. Tese (Doutorado) Institut d 'Informatique - Namur, Bélgica, 1993.
- De Azevedo, F. M.; Brasil, L. M.; Oliveira, R. C. L. de.; 2000. Redes Neurais com Aplicações em Controle e em Sistemas Especialistas. Florianópolis: Bookstore, p. 401, 2000.
- De Azevedo, F. M.; Ferrari, G. L.; Angeloni, M. N. M.; Argoud, F. I. M.; Algarve, A. S. Shell para Desenvolvimento de Sistemas Especialistas Fuzzy usando Prolog. In: Proceedings do III Congresso Latino Americano de Engenharia Biomédica, p. 919-992, 2004.
- De Carvalho, L. M. F. Análise dos Métodos de Aquisição do Conhecimento e Desenvolvimento de um Sistema Baseado na Metodologia Kads. 1995. Dissertação (Mestrado em Engenharia Elétrica e Informática Industrial), Centro Federal de Educação Tecnológica do Paraná, Curitiba, 1995.
- De Ré, A. M. Um modelo para determinar o perfil de empreendedores de sucesso utilizando uma ferramenta Neuro-Fuzzy. 2000. Tese (Doutorado), Universidade Federal de Santa Catarina, Florianópolis, 2000.
- Devore, J. L. Probabilidade e Estatística para Engenharia e Ciências. Editora Thomson Pioneira, SP, 2006.

- Duarte, L. F. D. *Da Vida Nervosa nas Classes Trabalhadoras Urbanas*. Rio de Janeiro: Jorge Zahar em co-edição com o CNPQ, 1988.
- Engel, J. Jr. Reply to of cabbages and kings: Some considerations on classifications, diagnostic schemes, semiology and concepts. *Epilepsia*, v. 44, p. 4-6, 2003.
- Engel, J. Psychogenic seizures. Disponível em: <<http://www.medlink.com>>. Acesso em: 02 Abr. 2006.
- Engel, P. M. *Ferramentas de IA: Representação do conhecimento em Sistemas Inteligentes Avançados*. Universidade Federal do Rio Grande do Sul, Instituto de informática, Porto Alegre, 2005.
- Fausett, L. *Fundamentals of Neural Networks - Architectures, Algorithms and Applications*. New Jersey: Prentice Hall, 1994.
- Feigenbaum, E. A.; Mccorduck, P. *The Fifth Generation: artificial intelligence and Japan's Computer Challenge to the Word*, 1983.
- Fernandes, A. M. da R. *Inteligência Artificial - Noções Gerais*. Florianópolis: Bookstore Livraria Ltda., 159 p., 2003.
- Fernandes, J. G.; Schimidt M. I.; Monte, T. L.; Tozzi, S.; Sander, J. W. A. S. Prevalence of epilepsy: the Porto Alegre study. *Epilepsia*, v. 33, supl. 3:132, 1992.
- FERNANDES, J. G. *Epidemiologia das crises epilépticas em Porto Alegre: um estudo populacional*. 1993. Porto Alegre. Tese (Doutorado), Universidade Federal do Rio Grande do Sul, RS, 1993.
- Fernandes, P. T.; Souza, R. J. De.; Li, L. M. Relatório do III Encontro Nacional de Associações e Grupos de Pacientes com Epilepsia. *Journal of Epilepsy and Clinical Neurophysiology*, v. 11, n. 2, p. 97-99, 2005.
- Fisher, R. F. Editor's introduction: Cabbages and kings in the classification of seizures and the epilepsies. *Epilepsia*, v. 44, p.1, 2003.
- Freeman, J. A.; Skapura, D. M. *Neural Networks: algorithms, applications and Programming Techniques*. New York: Addison-Wesley Publishing Company, 1991.
- Freud, S. *Histeria: primeiros artigos*. Tradução de José Luís Meurer, Rio de Janeiro: Imago Ed., 104 p., 1998.
- Fu, L. M. Rule Generation from Neural Networks. *IEEE Transactions on Systems, Man and Cybernetics*, v. 24, n. 8, p. 1114-1124, 1994.
- Fullér, R. *Introduction to Neuro-Fuzzy Systems*. New York, USA : Physica-Verlag, 2000.

- Garcia-Feijoo, J.; Carmona Suarez, E. J.; Gallardo, L. M. et al. Development of an automatic discrimination system for glaucomatous visual fields based on Neuro-Fuzzy nets. *Archivos de la Sociedad Española de Oftalmología*, v. 77, n. 12, p. 669-676, 2002.
- Gates, J. R. Classification and differential diagnosis. In: *Non Epileptic Events, Pseudoseizures*. New York: Syllabus of the American of Neurology, 1993.
- Gordon, J.; Shortliffe, E. H. The Dempster-Shafer Theory of Evidence. In *Rule-Based Systems - The MYCIN Experiments of the Stanford Heuristic Programming Project*, Eds. B. G. Buchanan and E. H. Shortliffe, Addison-Wesley Pub. Comp, USA, 1984.
- Gorzalczany, M. B. *Computational Intelligence Systems and Applications. Studies in fuzziness and soft computing*. New York, USA: Physica-Verlag, v. 86, 2002.
- Guazzelli, A. *Aprendizagem em Sistemas Híbridos*. 1994. Dissertação (Mestrado), Universidade Federal do Rio Grande do Sul, Porto Alegre, 1994.
- Guerreiro, C. A. M.; Guerreiro, M. M.; Cendes, F.; Cendes, I. L. *Epilepsia*. São Paulo: Lemos Editorial, 2000.
- Gupta, M. M.; Rao, D. H. On the Principles of Fuzzy Neural Networks. *Fuzzy Sets and Systems*, Canada: Elsevier Science, v. 61, p. 1-18, 1994_{a}.
- Gupta, M. M.; Rao, D. H. *Neuro-Control Systems: A Tutorial*. IEEE Press, Saskatoon, Canada, 1994_{b}.
- Guyton, A. C.; Hall, J. E. *Tratado de Fisiologia Médica*. Tradução da 11ª ed. Rio de Janeiro: Editora Elsevier, 2006.
- Hahne, M. N. M. A. *Metodologia de Aprendizado Baseado em Algoritmos Genéticos para Modelos Alternativos de Redes IAC*. 2005. Tese (Doutorado em Engenharia Elétrica) - Curso de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2005.
- Halgamuge, S. K.; Glesner, M. Fuzzy neural fusion techniques for industrial applications. In: *ACM Symposium on Applied Computing (SAC '94)* Phoenix, USA, March, 1994.
- Halgamuge, S. K.; Glesner, M. A Fuzzy neural approach for pattern classification with the generation of rules basedon supervised learning. In: *Neuro-Nines 92*, Nanterre, France, November, 1992.
- Haykin, S. *Redes Neurais: Princípios e Prática*. 2. ed. Porto Alegre : Bookman, 2001.
- Hebb, D. O. *The Organization of Behavior*. Ed. Wiley, 1949.
- Hebb, D. O. *Psicologia*. 2. ed. Philadelphia, USA: W. B. Saunders Company, 1979.
- Hirota, K.; Pedrycz, W. Knowledge-based networks in classification problems, *Fuzzy Sets and Systems*, v. 51, p. 1-27, 1992.

Hirota, K.; Pedrycz, W. OR/AND neuron in modeling fuzzy set connectives, IEEE Transactions on Fuzzy Systems, v. 2 p. 151-161, 1994.

Ilae. Commission on Classification and Terminology of International League Against Epilepsy. In: Proposal for Revised Clinical and Electrographic Classification of Epileptic Seizures, Epilepsia, v. 22, p. 489-501, 1981.

Ilae. Commission on Classification and Terminology of International League Against Epilepsy. In: Proposal for Revised Clinical and Electrographic Classification of Epilepsies and Epileptic Syndromes, Epilepsia, v. 30, p. 389-399, 1989.

Ilae. The 2001 Proposal for a Diagnostic Scheme. Disponível em: <<http://www.ilae-epilepsy.org/Visitors/Publications/AnnualReport2003/classification.cfm>>. Acesso em: 22 mar. 2001.

Ilae. A proposed Diagnostic Scheme for People with Epileptic Seizures and with Epilepsy: Report of the ILAE Task Force on Classification and Terminology. Disponível em: <<http://www.ilae.org>>. Acesso em: 05 abr. 2006.

Iaione, F. Detecção de Hipoglicemia Baseada na Análise e Classificação do Eletroencefalograma. 2003. Florianópolis. Tese (Doutorado em Engenharia Elétrica) - Curso de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Santa Catarina, SC, 2003.

Jang, J. -S. R. Adaptative-network-based fuzzy inference systems. IEEE Transactions on Systems, Man and Cybernetics, v. 23, n. 3, p. 665-658, 1993.

Jang, J. -S. R.; Sun, C. -T.; Mizutani, E. Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence. New Jersey : Prentice-Hall, 1997.

Javad, T.; Karim, M.; Majid, O. A fuzzy hibrid learning algorithm for radial basis function neural network with application in human face recognition. Pattern Recognition, Elsevier Science, v. 36, n. 5, May, p. 1187-1202, 2003.

Kartalopoulos, S. V. Understanding Neural Networks and Fuzzy Logic: Basic Concepts and Applications. New York: USA: IEEE Press, Inc, 1996.

Klir, G. J.; Folger, T. A. A Fuzzy sets uncertainty and information. Prentice Hall, 1988.

Klir, G.; Yuan, B. Fuzzy Sets and Fuzzy Logic: Theory and Applications. USA: Prentice Hall, 1995.

Koehler, C. Modelagem de Redes Bayesianas a partir de padrões identificados em Bases de Dados. 2003. Porto Alegre. Exame de Qualificação apresentado no Programa de Pós-Graduação em Ciência da Computação, Universidade Federal do Rio Grande do Sul, 2003.

Kovács, Z. L. Redes Neurais Artificiais - Fundamentos e Aplicações. São Paulo: Editora Acadêmica, 1996.

- Kröse, B. J. A.; Smagt, P. P. van der. *An Introduction to Neural Networks*. 7. ed. Amsterdam: The University of Amsterdam, 1995.
- Kwan, H. K. E Cai, T. *A Fuzzy Neural Network and its Application to Pattern Recognition*, IEEE Transactions on Fuzzy Systems, v. 3, p. 185-193, 1994.
- Lee, S. C.; Lee, E. T. Fuzzy sets and neural networks, *Journal of Cybernetics*, v. 4, p. 83-103, 1974.
- Likert, R. A. A technique for the development of attitude scales. *Educational and Psychological Measurement*, v. 12, p. 313-315, 1952.
- Li, R-P.; Mukaidono, M.; Turksen, I. B. A fuzzy neural network for pattern classification and feature selection. *Fuzzy Sets and Systems*, v. 130, p. 101-108, 2002.
- Lim, F. J.; Wai, R. J. Adaptive and fuzzy neural network sliding-mode controllers for motor-quick-return servomechanism. *Mechatronics*, v. 13, n. 5, p. 477-506, jun. 2003.
- Lim, J. S.; Wang, D.; Kim, Y.; Gupta, S. A neuro-fuzzy approach for diagnosis of antibody deficiency syndrome. Disponível em: <<http://www.sciencedirect.com>>. Acesso em: 10 nov. 2005.
- Lin, C. T.; Lee, C. S. G. Real-time supervised structure/parameter learning for fuzzy neural networks. In: *IEEE International Conference on Fuzzy Systems*, San Diego, USA, p. 1283-1291, 1992.
- Lüders, H.; Hajm, I.; Wyllie, E. Reply to of cabbages and kings: Some considerations on classifications, diagnostic schemes, semiology and concepts. *Epilepsia*, v. 44, p. 6-8, 2003.
- Luger, G. F. *Inteligência Artificial: Estruturas e Estratégias para a solução de Problemas Complexos*. Trad. Paulo Engel. 4. ed. Porto Alegre: Bookmann, 2004.
- Luger, G. F.; Stubblefield, W. A. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. Benjamim/Cummings, Redwood City, 2. ed. California, 1993.
- Mamdani, E. H. Application of fuzzy algorithm for control of simple dynamic plant. In: *Proceedings of IEEE Control and Science*, v. 121, n. 12, p. 1585-1588, 1974.
- Mcculloch, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115-133, 1943.
- Mcdermott, D.; DOYLE, J. Non-monotonic logic I. *Artificial Intelligence*, v. 13, n.1-2, p. 41-72, 1980.
- Melin, P.; Castillo, O. *Hybrid Intelligent Systems for Patern Recognition Using Soft Computing*. Studies in Fuzziness and Soft Computing, New York, Springer Berlin Heidelberg, v. 172, 2005.
- Minsky, M.; Papert, S. *Perceptrons: an introduction to computacional geometry*. MIT Press, Massachusetts, 1969.

Mit. Massachusetts Institute of Technology: The Artificial Intelligence Laboratory. Disponível em: <<http://www.ai.mit.edu/>>. Acesso em: 25 jan. 2007.

Mitra, S.; De, R. K.; Pal, S. K. Knowledge-Based Fuzzy MLP for Classification and Rule Generation. IEEE Transactions on Neural Networks, New York, v. 8, n. 6, p. 1338-1350, nov. 1997.

Mitra, S.; Pal, S. K. Fuzzy Multi-Layer Perceptron, Inferencing and Rule Generation. IEEE Transactions on Neural Networks, New York, v. 6, n. 1, p. 51-63, nov. 1995.

Mitra, S.; PAL, S. K. Multi-Layer Perceptron, Fuzzy Sets and Classification. IEEE Transactions on Neural Networks, New York, v. 3, n. 5, p. 683-697, set. 1992.

Mitra, S.; PAL, S. K. Logical Operation Based Fuzzy MLP for Classification and Rule Generation. Neural Networks, USA, v. 7, n. 2, p. 353-373, 1994.

Nassar, S. M. Estatística e Informática: um processo interativo entre duas ciências. Florianópolis, Universidade Federal de Santa Catarina, Departamento de Informática e Estatística, SC, 1998.

Nauck, D. A Fuzzy Perceptron as a Generic Model for Neuro-Fuzzy Approaches. In: Proc. Fuzzy-Systeme - 2nd GI-Workshop, Munich, Siemens Corporation, p. 91-99, 1994.

Nauck, D.; Kruse, R. Choosing Appropriate Neuro-Fuzzy Models. In Proceedings Second European Congress on Fuzzy and Intelligent Technologies, p. 552-557, 1994_{a}.

Nauck, D.; Kruse, R. NEFCON-I. An X-Windows based Simulator for Neural Fuzzy Controllers. In Proceedings IEEE Int. Conf. Neural Networks, p. 1638-1643, 1994_{b}.

Nauck, D. Beyond Neuro-Fuzzy: Perspectives and Directions. Third European Congress on Intelligent Techniques and Soft Computing, August, p. 1159-1164, 1995.

Nauck, D.; Kruse, R. A NEFCLASS-Fuzzy Approach for the Classification of Data. In: Proc.of the 1995 ACM Symposium on Applied Computing. New York, p. 461-465, fev. 1995.

Nauck, D.; Kruse, R. Designing Neuro-Fuzzy Systems through backpropagation. In: Witold Pedrycz, editor, Fuzzy Modelling: Paradigms and Practice. Kluwer, p. 203-228, 1996.

Nauck, D.; Kruse, R. What are Neuro-Fuzzy Classifiers? In: Seventh International Fuzzy Systems Association World Congress, v. IV, p. 228-233, 1997_{a}.

Nauck, D.; Kruse, R. A Neuro-Fuzzy method to learn fuzzy classification rules from data. Fuzzy Sets and Systems, v. 89, p. 277-288, 1997_{b}.

Nauck, D.; Kruse, R. Neuro-fuzzy Systems for Function Approximation. Fuzzy Sets and Systems, v. 101, p. 261-271, 1999.

- Ojala, T. Neuro-Fuzzy Systems in Control. 1994. Finland. Thesis (Master of Science). Tampere University of Technology, 1994.
- Osório, F. S.; Amy, B.; Decloedt, L. Rule-Out: Um novo método de extração incremental de regras à partir de redes neurais construtivas do tipo KBANN. SBRN'98 - Simpósio Brasileiro de Redes Neurais, Belo Horizonte, 1998.
- Osório, F. S. O.; Vieira, R. Sistemas Híbridos Inteligentes. XIX Congresso da Sociedade Brasileira da Computação - Encontro Nacional de Inteligência Artificial. Disponível em <<http://www.inf.unisinos.br/~osorio/enia99/>>. Acesso em: 20 fev. 2007.
- Pal, S. K.; Mitra, S. Multilayer Perceptron, Fuzzy Sets and Classification. IEEE Transactions on Neural Networks, New York, v. 3, n. 5, p. 683-697, set. 1992.
- Parker, D.; 1982. Learning Logic. Invention Report, Stanford University, File 1, Office of Technology Licensing, Stanford, California, p. S81-64.
- Passold, F. Controle Neural de Posição de Força em Manipuladores Robóticos. 2003. Tese (Doutorado em Engenharia Elétrica), Universidade Federal de Santa Catarina, Florianópolis, 2003.
- Pedrycz, W. A Referencial Scheme of Fuzzy Decision Making and Its Neural Network Structure. IEEE Transactions on Systems, v. 21., n. 6, p. 1593-1604, 1991.
- Pedrycz, W. Fuzzy Neural Networks and Neurocomputations. Fuzzy Sets and Systems, v. 56, p. 1-28, 1993.
- Pedrycz, W.; Rocha, A. F. Fuzzy-Set Based Models of Neurons and Knowledge-Based Networks. IEEE Transactions on Fuzzy Systems, v. 1., n. 4, p. 254-266, 1993.
- Pedrycz, W. Heterogeneous fuzzy Logic Networks: Fundamentals and Development Studies. IEEE Transactions on Neural Networks, v. 15, n. 6, p. 1466-1481, nov. 2004.
- Peil, N. de C. Reconhecimento de Voz: Uma Abordagem Utilizando Lógica Difusa. 1998. Dissertação (Mestrado em Ciência da Computação), Universidade Federal de Santa Catarina, Florianópolis, 1998.
- Raittz, R. T. Fan 2002: Um Modelo Neuro-Fuzzy para Reconhecimento de Padrões. 2002. Tese (Doutorado), Universidade Federal de Santa Catarina, Florianópolis, 2002.
- Ramani, V. Review of psychiatric treatment strategies in non epileptic seizures. In: Rowan A.J.; Gates, J.R..Non Epileptic Seizures. Butterworth-Heinemann, p. 259-268, 1993.
- Raposo, R. de C. T. Uma Ferramenta para Análise Fundamentalista de Investimentos em Empresas Utilizando Sistemas Neuro-Nebulosos. 2001. Dissertação (Mestrado), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2001.

- Rezende, S. O.; Evsukokk, A. G.; Garcia, A. C. B. et al. *Sistemas Inteligentes - Fundamentos e Aplicações*. São Paulo: Editora Manole, 2003.
- Rosenblatt, F. The Perceptron: A probabilistic model for information storage and organization in the brain. *Psychol. Rev.* v. 65, p. 386-408, 1958.
- Rumelhart, D. E.; McClelland, J. L. *Explorations in Paralled Distributed Processing*. Foundations The MIT Press, v. 1, 1986_a.
- Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagation errors. *Nature*, v. 323, n. 9, p. 451-453, 1986_b.
- Russell, S.; Norvig, P. *Artificial Intelligence: A Modern Approach*. 2. ed. New Jersey: Prentice Hall, 1995.
- Russell, S.; Norvig, P. *Inteligência Artificial*. Tradução da 2. ed. Rio de Janeiro: Elsevier, 2004.
- Rutkowska, D. *Neuro-Fuzzy Architectures and Hybrid Learning*. Studies in fuzziness and soft computing, New York: Physica-Verlag Heidelberg, v. 85, 2002.
- Santo, J. L. E.; Maineri, N.; Portuguesez, M. W. Epilepsia e Crises Não-Epilépticas. *Journal of Epilepsy and Clinical Neurophysiology*, v. 4, p. 29-33, 2004.
- Serhatlioglu, S.; Bozgeyik, Z.; Ozkan, Y.; Hardalac, F.; Güler, I. Neuro-fuzzy classification of the effect of diabetes mellitus on carotid artery. *Journal Medical Systems*. v. 27, n. 5, p. 457-464, 2003.
- Serhatlioglu, S.; Hardalaç, F.; Kiris, A.; Ozdemir, H.; Yilmaz, T.; Güler, I. A neuro-fuzzy classification system for the effects of diabetes mellitus on ophtalmic artery. *Journal Medical Systems*. v. 28, n. 2, p. 167-176, 2004.
- Shafer, G. *Mathematical Theory of Evidence*. Princeton, NJ, Princeton University Press, 1976.
- Sigart. Special Interest Group in Artificial Intelligence. Disponível em: <<http://www.acm.org/sigart>>. Acesso em: 16 jan. 2006.
- Silva, P. *As Bases Farmacológicas da Terapêutica*. Trad. Goodman & Gilman, Rio de Janeiro: Guanabara Koogan, 1987.
- Simpson, P. K. Fuzzy Min-Max Neural Networks - Part 1: Classification. *IEEE Transactions on Neural Networks*, v. 3, n. 5, p. 776-786, 1992.
- Smithson, M. *Ignorance and Uncertainty: emerging paradigms*. Springer-Verlag. New York, 393 p., 1989.
- Souza, F. J. *Modelos Neuro-Fuzzy Hierárquicos*. 1999. Tese (Doutorado), Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 1999.

- Sugeno, M.; Kang, G. T. Structure Identification of Fuzzy Model. *Fuzzy Sets and Systems*, North-Holland, v. 28, p. 15-33, 1988.
- Takagi, H.; Hayashi, I. Constructions of fuzzy inference rules by ndf and ndfl. *International Journal of Approximate Reasoning*, v. 6, p. 241-266, 1992.
- Takagi, T.; Sugeno, M. Fuzzy identification of systems and its applications to modelling and control. *IEEE Transactions On Systems, Man and Cybernetics SMC*, v. 15, n. 1, p. 116-132, 1985.
- Tierney, L. M. Jr.; Mcphee, S. J.; Papadakis, M. A. *Diagnóstico e Tratamento*. São Paulo: Atheneu Editora, 2004.
- Tsukamoto, Y. An approach to fuzzy reasoning method. In Madan M. Gupta, Rammohan K. Ragade, and Ronald R. Yager (Eds), *Advances in fuzzy set theory and applications*, p. 137-149. North Holland, Amsterdam, 1979.
- Velasco, M. M. B. R.; Pacheco, M. A. C. Núcleo de Pesquisa em Inteligência Computacional Aplicada. PUC-RIO. Disponível em <<http://www.ica.ele.puc-rio.br>> Acesso em: 14 jul. 2006.
- Velasco, M. M. B. R.; Pacheco, M. A. C.; Neto, L. S. R.; Souza, F. J. de. Electric load forecasting: evaluating the novel hierarchical neuro-fuzzy BSP model. *International Journal of Electrical Power & Energy Systems*. Published by Elsevier Science Ltda, v. 26, n. 2, p. 131-142, 2004.
- Vellido, A.; Cerquides, J. Neural Networks For B2C E-Commerce Analysis: some elements of best practice. In: *Proceedings of Fourth International Conference on Enterprise Information Systems*, April, Ciudad Real, Spain, p. 439-445, 2002.
- Viali, L. Um modelo neuro-difuso de multiagentes para apoio ao ensino de disciplinas de ciências exatas. 1999. Tese (Doutorado) Universidade Federal de Santa Catarina, Florianópolis, 1999.
- Vianna, G. K. *Sistemas Inteligentes de Arquitetura Híbrida Neuro-Difusa para Reconhecimento de Padrões*. 2000. Rio de Janeiro. Dissertação (Mestrado), Universidade Federal do Rio de Janeiro, RJ, 2000.
- Yu, W.; Lai, C.; Lee, W. A WICE approach to real-time construction cost estimation. *Automation in Construction*, v. 15, p. 12-19, 2006.
- Wagner, A. *Extração de Conhecimento a partir de Redes Neurais aplicada ao Problema da Cinemática Inversa na Robótica*. 2003. Dissertação (Mestrado), Universidade do Vale do Rio dos Sinos, São Leopoldo, 2003.
- Werbos, P. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. Ph.D. Thesis, Harvard University, Cambridge, Massachusetts, 1974.
- Widrow, B.; Hoff, M. E. Adaptive switching circuits. In: *1960 IRE WESCON Convention Record*, p. 96-104, 1960.

Widrow, B. Generation and information storage in networks of adaline neurons. In: Self-Organizing Systems, p. 435-461. Spartan Books, Washington, D. C., M. Yovits and G. Gacobi and G. Goldstein (Eds), 1962.

Wolf, P. Of cabbages and kings: Some considerations on classifications, diagnostic schemes, semiology and concepts. *Epilepsia*, v. 44, p. 1-4, 2003.

Who. The CID-10 classification of mental and behavioural disorders: clinical descriptions and diagnostic guidelines. Geneve: Division of Mental Health, 1993.

Zadeh, L. A. Fuzzy Sets, *Information and Control*, v. 8, n. 3, p. 338-353, 1965.

Zadeh, L. A. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man and Cybernetics*, v. SMD-3, n. 1, p. 28-44, 1973.

Zadeh, L. A. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, v. 1, p. 3-28, 1978.

Zadeh, L. A. The Berkeley Initiative in Soft Computing (BISC). Disponível em: <<http://www.cs.berkeley.edu/~zadeh/>>. Acesso em: 16 jan. 2006.

ANEXOS

ANEXO 1 - Formulário para Coleta de Dados do Diagnóstico em ENEs
Evento Não Epiléptico (ENE)

Nome do Paciente:

Endereço: Estado: Telefone:

Idade do paciente: Idade Inicial da Crise:

Queixa Principal:

Relato da Crise:

<i>Sintomas:</i>	Às Vezes	Frequentemente	Sempre
1- Quadros Clássicos Conversão:			
Espasmos da Faringe/Perturbações Digestivas	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Contraturas/Dores no corpo/Choro	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Paralisia:			
Afasia (perda voz)/Afonía (escuta)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Astasia (andar)/Abasia (ficar de pé)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 – Distúrbios:			
Sensibilidade (anestesia)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Atividade Sensorial	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 – Fatores Emocionais/Físicos:			
Alucinações e/ou Delírios/Palpitações/Nervos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Falta de ar/Enxaqueca/Insônia/Falta de apetite/	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tontura/Tremores/Irritação/Depressão/Stresse/	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Esgotamento/Tristeza/Fúria/Raiva/Brigas/Desajustes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 – Transtornos:			
Pânico/Comportamental/Fugas/Conversivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Despersonalização/Dissociativo/Somatização/	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Hipocondria/Simulação/Explosivo	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Outros	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 – Síncopes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6 – Distúrbios:			
Sono (Terror Noturno/Sonambulismo/Bruxismo)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Movimentos Arritmicos (Tiques em vários músculos)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Endócrinos (Glândulas de secreção internas)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Tóxicos/ Crise Febril	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

ANEXO 2 - Coleta de Dados do Diagnóstico em Eventos Epilépticos

	Às Vezes	Freqüentemente	Sempre
7 - Sinais e Sintomas Parciais:			
Sintomas Motores	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Somatosensoriais	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Autonômicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Físicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Houve Prejuízo da Consciência ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sentiu cheiros/gosto estranhos ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 - Durante o desmaio “viu” alguma coisa ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 - Durante o desmaio “ouviu” alguma coisa ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10 - Durante o desmaio mordeu a língua ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Durante o desmaio urinou-se ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11 - Durante o desmaio machucou-se/feriu-se ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12 - Tempo de Duração – Sua crise durou:			
<input type="checkbox"/> Até 1 (um) minuto ?	<input type="checkbox"/> Até 5 (cinco) minutos ?	<input type="checkbox"/> Até 30 (trinta) minutos ou mais ?	
13 – Eventos Pós-Críticos:			
Após a crise houve confusão mental ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Após a crise houve sonolência ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Após a crise houve paralisias ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Após a crise houve vômito/náuseas/tonturas ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Após a crise houve dor de cabeça e/ou dores corpo?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Convulsões Recidivantes ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14 – Sinais e Sintomas Generalizados:			
Sintomas Mioclônicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas de Ausência	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Tônicos-Clônicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Clônicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Tônicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Sintomas Atônicos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Horário de Ocorrência da Crise?	Dia <input type="checkbox"/>	Noite <input type="checkbox"/>	
Paciente usa Álcool ou Drogas?	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	
Existe na Família casos Epilépticos ?	Sim <input type="checkbox"/>	Não <input type="checkbox"/>	
Freqüência da Crise: <input type="checkbox"/>			
(nº crises)			

ANEXO 3 - Termo de Consentimento Livre e Esclarecido (TCLE)

Meu nome é LUCIMAR M. FOSSATTI DE CARVALHO e estou desenvolvendo a pesquisa **MODELOS DE APRENDIZAGEM PARA SISTEMAS NEURO-DIFUSOS**, com o objetivo de criar e avaliar modelos de aprendizagem utilizados nos Sistemas de Apoio ao Diagnóstico Médico, em especial aplicados ao tratamento da imprecisão envolvendo os Eventos Epilépticos e os Eventos Não Epilépticos. Os modelos de aprendizagem, após avaliação e validação, serão utilizados na minha tese de doutorado. O estudo é necessário porque a pesquisa precisa fazer o cadastramento dos sintomas apresentados pelos pacientes, para posteriormente, aplicar as técnicas de tratamento da imprecisão, tendo como objetivo obter melhores percentuais de acertos do sistema. Pretende-se, ao final da pesquisa, que o sistema seja disponibilizado nos ambulatórios disponíveis da rede pública.

Se você estiver de acordo em participar, posso garantir que as informações fornecidas serão confidenciais e serão apenas utilizadas nesta pesquisa.

Orientanda: Lucimar M. Fossatti de Carvalho (RG 100 975 6535 SSP/RS)

Orientador: Prof^o Dr. Fernando Mendes de Azevedo (UFSC/SC)

Co-orientadora: Prof^a Dra. Sílvia Modesto Nassar (UFSC/SC)

Consentimento Pós-Informação

Eu, < informar o nome do paciente ou responsável > , fui esclarecido sobre a pesquisa **MODELOS DE APRENDIZAGEM PARA SISTEMAS NEURO-DIFUSOS** e, ao responder estes dois questionários, concordo que meus dados sejam utilizados apenas para pesquisa.

Assinatura do paciente ou responsável

Instruções

Preencher as colunas dos questionários de acordo com a intensidade dos sintomas apresentados, selecionando uma das opções oferecidas juntamente com o total de sintomas ocorridos no mês:

Opção_1: Às Vezes

Opção_2: Frequentemente

Opção_3: Sempre

Todos os itens devem ser preenchidos, assinalando apenas uma **única** opção. O objetivo da aplicação deste questionário é quantificar os sintomas ocorridos no paciente. A sinceridade e espontaneidade são extremamente importantes para a validade da pesquisa.



RESOLUÇÃO

O Comitê de Ética em Pesquisa-UPF, que é reconhecido pelo CONEP, na reunião do dia 21/01/2004, analisou o projeto:

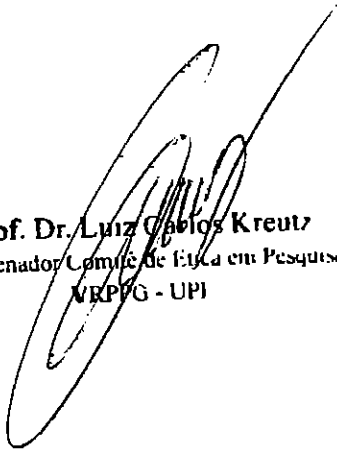
PROJETO nº: 223/2003

TÍTULO: Sistema Neuro-Difuso: Modelos de Aprendizagem

PESQUISADOR RESPONSÁVEL: Lucimar M. Fossatti de Carvalho

Este projeto foi analisado e **APROVADO**, por atender as normas e diretrizes da Resolução CNS 196/96 do Conep.

Passo Fundo, 05 de março de 2004.



Prof. Dr. Luiz Carlos Kreutz
Coordenador Comitê de Ética em Pesquisa
VRPPG - UPI

APÊNDICES

APÊNDICE 1 - Base de Treinamento

00|00|01|00|00|00|00|00|01|00|05|02,0|01|05|100
00|00|01|00|00|00|01|00|00|00|00|05|01,5|00|04|100
00|00|00|00|00|00|00|00|00|00|00|05|01,0|02|03|100
05|01|02|00|01|00|00|00|00|00|00|05|02,0|00|15|010
00|01|04|00|00|00|00|00|00|00|00|08|00,0|00|04|010
00|00|00|00|00|00|00|00|00|00|00|05|02,5|02|04|100
00|00|00|00|01|00|00|00|00|01|00|05|03,5|04|04|100
05|08|00|00|10|12|00|00|00|00|00|20|01,0|00|12|010
02|05|02|00|00|00|00|00|00|00|00|10|03,0|00|15|010
00|00|00|00|00|00|01|00|00|00|01|08|00,0|03|03|100
00|00|00|00|00|00|00|00|00|00|02|05|03,0|10|10|100
01|00|04|01|00|02|00|00|00|00|00|05|00,0|00|04|010
00|00|01|00|02|00|00|00|00|00|00|20|00,0|00|04|010
00|00|00|00|00|00|01|00|00|00|00|05|01,0|01|01|100
00|00|00|00|00|00|01|00|00|00|00|05|01,0|01|01|100
00|00|00|00|00|00|01|00|00|00|00|05|02,0|01|02|100
00|00|10|00|01|00|00|00|00|00|00|30|00,0|00|20|010
01|01|01|01|00|01|00|00|00|00|00|30|00,0|00|01|010
00|01|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
00|04|03|02|00|00|00|00|00|00|00|30|00,0|00|04|010
00|00|00|00|00|00|00|00|00|00|00|01|00,0|01|01|100
01|00|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
05|04|01|03|00|00|00|00|00|00|00|30|00,0|00|05|010
00|01|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|00|01|00|00|00|00|01|00,0|00|01|100
04|03|05|01|00|00|00|00|00|00|00|30|00,0|00|05|010
01|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|00|01|00|00|00|00|01|01,0|01|01|100
00|00|00|00|00|00|04|00|00|05|04|30|03,0|03|05|100
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|03|00|00|03|04|30|03,0|03|04|100
00|00|00|00|00|00|00|00|00|00|00|01|06,0|00|06|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|03|00|00|02|03|30|03,0|02|03|100
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|01|00|00|00|00|01|00,0|01|01|100
03|04|02|01|00|00|00|00|00|00|00|05|00,0|00|04|010
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|00|00|00|00|00|00|00|05|04,0|04|05|100
00|00|00|00|00|00|01|00|00|00|00|01|00,0|00|01|100
00|00|01|00|00|00|00|00|00|00|00|05|02,0|02|02|100
05|04|06|00|00|00|00|00|00|00|00|05|00,0|00|06|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|00|00|00|00|00|00|00|00|30|00,0|00|01|010
00|00|00|00|00|00|03|00|00|00|00|05|05,0|06|06|100
00|01|01|01|00|00|00|00|00|00|00|20|00,0|00|01|010

00|00|00|00|00|00|00|00|00|00|00|05|02,0|02|02|100
05|03|07|00|00|00|00|00|00|00|00|05|00,0|00|08|010
01|00|01|00|00|01|01|00|00|00|00|01|00,0|01|01|100
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|02|03|05|02,0|03|03|100
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|01|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|03|04|05|00|00|00|00|00|00|05|00,0|00|05|010
00|00|00|00|00|00|01|00|00|00|00|01|01,0|01|01|100
00|00|00|00|00|00|00|00|00|00|00|01|00,0|01|01|100
00|00|00|00|00|00|03|00|00|00|00|05|02,0|03|04|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
03|02|01|03|00|00|00|00|00|00|00|05|00,0|00|03|010
00|00|00|00|00|00|01|00|00|01|00|01|01,0|00|01|100
01|01|00|00|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|00|00|00|00|00|00|00|01|02|05|02,0|03|04|100
00|00|00|00|00|00|00|00|00|01|01|01,0|01|01|100
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|02|03|00|00|00|00|00|00|05|00,0|00|03|010
00|00|01|00|00|00|00|00|00|00|00|01|01,0|02|02|100
00|00|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|03|02|05|03,0|04|05|100
00|00|00|00|00|00|00|00|00|01|00|01|01,0|00|01|100
00|01|01|01|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|01|06|00|00|00|00|00|00|00|05|00,0|00|07|010
00|00|00|00|00|00|00|00|00|00|00|01|02,0|02|03|100
01|00|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
02|00|03|00|00|00|03|00|00|00|00|05|00,0|03|04|010
01|01|00|00|00|01|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|01|00|00|00|00|01|00,0|00|02|100
01|02|01|02|00|00|00|00|00|00|03|05|03,0|00|04|100
01|01|01|01|01|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|00|00|01|03,0|03|03|100
02|03|01|00|00|00|00|00|00|02|00|05|00,0|03|03|010
01|00|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|01|01|01|00|00|00|00|05|00,0|00|01|010
03|02|02|00|00|00|00|00|00|00|00|05|02,0|04|05|100
00|00|01|00|00|00|00|00|00|00|00|01|01,0|01|01|100
01|00|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|02|03|00|00|00|00|00|02|00|05|02,0|03|06|010
01|00|00|01|00|01|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|00|00|01|00,0|01|01|100
01|00|03|02|00|00|00|00|00|00|02|05|02,0|03|03|100
01|00|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|00|00|01|01,0|00|02|100
01|04|04|00|00|00|00|00|00|00|03|05|04,0|00|04|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|00|00|00|00|00|00|00|01|00,0|00|01|100
01|02|03|00|00|00|00|00|00|04|00|05|00,0|04|05|100

01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|01|00|00|00|00|00|00|05|00,0|00|01|010
08|05|01|04|00|00|00|00|00|00|03|05|00,0|04|08|010
00|00|00|00|00|00|00|00|00|00|01|01,0|01|01|100
00|00|01|01|00|01|00|00|00|00|00|05|00,0|00|01|010
01|00|02|03|00|00|00|00|00|00|02|05|02,0|04|05|100
00|00|00|00|00|00|00|00|00|01|00|05|01,0|01|01|100
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|03|02|00|00|04|00|00|00|02|05|03,0|00|06|010
00|00|00|00|00|00|00|00|00|00|00|30|01,0|01|01|100
00|00|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
03|00|01|04|00|00|00|00|00|00|02|05|04,0|03|04|100
00|00|00|00|00|00|00|00|00|00|00|05|01,0|01|01|100
01|01|01|00|01|00|00|00|00|00|00|05|00,0|00|01|010
01|00|02|00|03|00|00|00|00|00|03|05|02,0|02|04|010
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|03|00|02|00|00|00|00|00|02|05|02,0|01|03|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|02|02|00|00|00|00|00|01|00|05|01,0|02|02|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|02|04|03|00|00|01|00|00|00|00|05|04,0|03|05|010

Obs: Os quatorze primeiros registros deste formulário representam a codificação dos sintomas apresentados pelos pacientes, cadastrados na base de treinamento. O registro nº quinze representa a frequência da crise e o registro dezesseis, a saída desejada: “100” para os eventos epiléticos e “010” eventos não epiléticos.

APÊNDICE 2 - Base de Teste

00|00|00|00|00|00|00|00|00|00|00|01|01,5|03|03|100
01|00|00|00|00|01|00|00|00|00|00|01|02,0|00|02|100
01|01|02|00|02|00|00|00|00|00|00|30|00,0|00|02|010
00|00|00|00|00|00|30|00|00|00|00|01|00,0|00|30|100
00|00|01|00|01|00|03|00|00|00|00|01|01,5|01|04|100
00|00|00|00|00|00|00|00|00|00|05|06|05|0,10|06|10|100
00|00|08|00|01|05|00|00|00|00|00|03|00,0|00|15|010
00|02|00|00|00|00|00|00|00|00|05|00|05|03,0|02|06|100
00|00|01|00|00|00|01|00|00|00|00|03|01,0|00|02|100
00|00|00|00|00|00|00|00|00|01|00|03|01,0|01|01|100
00|01|01|00|02|00|00|00|00|00|00|05|00,0|00|02|010
00|01|00|00|00|00|00|00|00|02|00|03|01,0|02|02|100
00|00|01|00|00|00|01|00|00|00|00|05|02,0|00|02|100
00|00|00|00|00|00|02|00|00|00|00|04|02,0|02|02|100
00|01|02|00|02|01|00|02|00|00|00|30|00,0|00|03|010
00|00|01|00|02|00|15|00|00|00|00|01|0,15|00|15|100
02|00|01|02|02|01|00|00|00|00|00|30|00,0|00|03|010
00|00|00|00|00|00|00|00|00|00|01|03,0|03|03|100
00|00|02|00|00|00|02|00|00|00|00|01|02,0|00|02|100
00|00|02|00|02|00|00|02|02|00|00|05|00,0|00|02|010
00|00|00|00|00|00|01|00|00|00|00|01|01,0|00|01|100
00|00|00|00|00|01|00|00|00|01|00|05|01,0|00|03|100
00|00|01|00|01|00|00|00|00|00|00|05|01,0|01|01|010
00|01|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|00|01|01,0|01|02|100
00|01|01|01|00|00|00|00|00|00|00|30|01,0|01|01|010
00|00|00|00|00|00|01|00|00|00|00|01|01,0|00|01|100
00|00|00|00|00|00|00|00|02|02|05|04,0|00|04|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|01|01|00|00|00|00|00|05|01,0|00|01|010
00|00|01|00|00|00|02|00|00|02|02|05|01,5|02|03|100
00|00|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|01|00|00|00|00|00|00|01|00,0|01|01|100
00|00|00|00|00|00|00|00|00|03|02|01|02,0|04|04|100
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|02|03|04|00|00|00|00|00|00|00|15|00,0|00|05|010
00|00|01|00|00|00|01|00|00|00|00|01|01,0|00|01|100
00|00|00|00|00|00|01|00|00|00|00|01|01,0|01|01|100
00|00|00|00|00|00|00|00|00|00|02|01|03,0|01|04|100
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|00|00|00|01|00|00|00|00|05|00,0|00|01|010
02|00|03|06|00|00|00|00|00|00|00|20|00,0|00|07|010
00|00|00|00|00|00|03|00|00|00|00|01|00,0|03|06|100

00|00|01|00|00|00|01|00|00|00|00|01|00,0|01|01|100
00|02|03|04|00|00|00|00|00|00|00|10|00,0|00|04|010
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|01|00|00|01|00|00|00|01|05|00,0|00|01|100
00|00|00|00|00|00|00|00|00|01|02|01|02,0|03|06|100
00|00|00|00|00|00|03|00|00|00|00|01|00,0|03|03|100
00|01|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
02|00|01|02|03|00|00|00|00|00|00|01|00,0|00|03|010
01|00|02|00|00|00|00|00|00|00|00|05|00,0|00|02|010
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|02|03|01|02,0|03|05|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|00|00|00|00|00|00|00|00|05|00,0|00|02|010
02|00|01|03|05|00|00|00|00|00|00|01|00,0|00|05|010
00|01|00|01|00|01|01|00|00|00|00|30|00,0|00|01|010
00|00|00|00|00|00|01|00|00|00|00|01|00,0|01|01|100
00|00|00|00|00|00|03|00|00|00|00|01|03,0|04|04|100
00|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
04|03|01|00|00|00|00|00|00|00|00|01|02,0|03|04|010
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
03|00|01|04|00|00|00|00|00|00|03|01|03,0|01|06|100
00|00|01|00|00|00|01|00|00|00|00|01|01,0|01|01|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|03|01|00|00|04|00|00|00|00|03|01|04,0|01|04|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
01|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
03|00|01|02|00|00|02|00|00|00|00|01|02,0|01|03|100
00|01|01|00|00|00|00|00|00|00|00|30|00,0|00|01|010
00|00|00|00|00|00|01|00|00|00|00|01|01,0|00|01|100
02|04|03|00|00|00|00|00|00|00|03|01|00,0|04|06|010
00|00|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
01|00|01|00|00|01|00|00|00|00|00|30|00,0|00|01|010
02|03|01|00|00|00|00|00|00|00|00|05|02,0|03|03|100
01|01|01|00|00|01|00|00|00|00|00|30|00,0|00|01|010
01|01|01|00|00|00|00|00|00|00|00|30|00,0|00|01|010
04|02|01|00|00|00|04|00|00|00|00|05|03,0|02|05|010
01|00|01|01|00|00|01|00|00|00|00|30|01,0|00|01|010
01|01|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
02|02|01|00|00|00|00|00|00|00|00|05|02,0|02|02|100
00|00|00|00|00|00|00|00|00|00|00|01|00,0|01|01|100
01|01|01|00|00|00|01|00|00|00|00|05|01,0|00|01|010
02|04|05|00|00|00|00|00|00|03|00|05|00,0|05|06|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
02|03|01|00|00|00|00|00|00|00|03|01|03,0|01|04|100

01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|00|00|00|00|00|01|00,0|01|01|100
05|00|01|03|00|00|00|00|00|00|03|10|02,0|04|06|010
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
01|00|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
03|03|01|00|00|00|00|00|00|00|00|01|03,0|03|04|100
01|01|01|00|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|00|00|00|00|00|00|01|00|01|00,0|00|01|100
02|02|01|03|00|00|00|00|00|00|02|01|02,0|03|03|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|00|01|01|00|00|00|00|00|00|00|30|00,0|00|01|010
03|04|01|00|00|00|03|00|00|00|00|01|01,0|03|04|100
01|01|01|00|00|01|00|00|00|00|00|05|00,0|00|01|010
00|00|00|00|00|00|02|00|00|00|00|01|00,0|02|02|100
03|02|01|00|00|00|00|00|00|00|02|10|02,0|01|03|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|01|01|01|00|00|00|00|00|00|00|05|00,0|00|01|010
00|02|04|03|00|00|01|00|00|00|00|05|04,0|03|05|010
00|00|03|03|00|00|00|00|00|02|00|05|02,0|02|06|010
00|00|00|00|00|00|00|00|00|00|01|05|09,0|07|09|100
02|00|02|00|00|00|00|00|00|00|00|05|02,0|02|02|010
00|00|00|00|00|00|00|00|00|01|00|05|02,0|02|03|100
01|00|01|01|00|00|00|00|00|00|02|05|02,0|00|02|010
00|00|00|00|00|00|00|00|00|00|00|05|07,0|12|12|100
00|00|00|00|00|00|00|00|00|00|00|05|30,0|30|30|100
00|00|04|04|00|00|00|00|00|00|00|05|03,0|03|04|010
00|00|00|00|00|00|00|00|00|00|00|01|30,0|30|30|100
00|00|00|00|00|00|02|00|00|00|00|05|02,0|02|02|100
00|00|02|02|00|00|00|00|00|00|00|01|01,0|01|02|010
00|00|00|00|00|00|00|00|00|00|00|05|01,0|01|01|100
02|00|03|03|00|00|00|00|00|00|02|01|00,0|02|06|010

Obs: Os quatorze primeiros registros deste formulário representam a codificação dos sintomas apresentados pelos pacientes, cadastrados na base de treinamento. O registro nº quinze representa a frequência da crise e o registro dezesseis, a saída desejada: “100” para os eventos epiléticos e “010” eventos não epiléticos.

APÊNDICE 3 – Variáveis de Entrada

1 - Quadros Clássicos de Conversão

Os quadros clássicos de conversão envolvem os espasmos da faringe, perturbações digestivas, contraturas, dores no corpo, choro e paralisias, incluindo afasia, afonia, astasia e abasia.

Os espasmos da faringe são relacionados à sensação de bola na garganta apresentada pelo paciente em relação a sua faringe, ou seja, como se uma bola estivesse subindo do epigástrico para a garganta. Uma contratura refere-se ao ato ou efeito de contrair-se. Nas contraturas o aparelho muscular reage a pequenos estímulos, ocorrendo num músculo ou num grupo muscular; apresentam uma persistência incomum e não relaxam com o sono. As paralisias podem afetar todas as partes do corpo ou segmentos do mesmo, como mão, ombro, cotovelo; afasia, relacionam-se com a perda da palavra falada, escrita ou mímica, por alteração dos centros nervosos; afonia, perda da capacidade de audição; astasia, com a incapacidade de andar do paciente, e abasia, com a incapacidade do paciente de ficar de pé.

2 - Distúrbios

Os distúrbios referem-se aos distúrbios da sensibilidade, relacionados com a anestesia, que consiste na privação da sensibilidade, afetando a pele, as membranas mucosas, ossos, músculos, nervos, órgãos dos sentidos e os intestinos. A anestesia pode ser total ou atingir apenas a sensação de dor (analgesia). Os distúrbios de sensibilidade podem afetar todos os órgãos dos sentidos.

3 - Fatores Emocionais

Referem-se a uma série de sintomas, entre os quais estão alucinações e/ou delírios, palpitações, falta de ar, enxaqueca, insônia, falta de apetite, tontura, tremores, irritação, depressão, estresse, esgotamento, tristeza, fúria, raiva e desajustes no ambiente.

4 - Transtornos

Os transtornos mais comuns são os de pânico, comportamental, fugas, conversivo, despersonalização, dissociativo, somatização, hipocondria, simulação e explosivo e outros.

5 - Síncopes

As síncopes estão relacionadas com breves perdas da consciência, geralmente com rápida recuperação, ocorrem em razão de múltiplos fatores, desde um simples susto (ansiedade, tensão emocional) até um quadro encefalítico.

6 – Outros Distúrbios

Esses distúrbios caracterizam-se por distúrbios do sono, incluindo o terror noturno, sonambulismo, bruxismo. Movimentos arrítmicos (tiques em vários músculos), endócrinos, incluindo as glândulas de secreção internas, tóxicos e crise febril.

7 - Sinais e Sintomas Parciais

Caracterizados por sintomas motores, somatossensoriais, autonômicos e físicos.

8 - Viu Alguma Coisa

Este sintoma relata se o paciente, durante o desmaio, viu alguma pessoa e/ou testemunha.

9 - Ouvia Alguma Coisa

Este sintoma relata quando o paciente, durante o desmaio, ouviu ou não algo de alguma testemunha próxima.

10 - Mordeu a Língua

Na classificação dos EPs, o ato do paciente morder a língua influencia na definição do diagnóstico. O sintoma refere-se ao ato do paciente ter mordido ou não a língua durante um EE.

11 - Machucou-se e/ou Feriu-se

Ato ou efeito do paciente, durante um EE, de machucar-se ou ferir-se.

12 – Tempo de Duração

Observar o tempo de duração da crise: até um minuto; até cinco minutos ou até 30 minutos ou mais.

13 - Eventos Pós-Críticos

Neste item são incluídos os principais sintomas que acometem os pacientes após uma crise convulsiva: confusão mental, sonolência, paralisias, vômito, náuseas ou tonturas, dor de cabeça ou dores no corpo.

14 - Sinais e Sintomas Generalizados

Caracterizados por sintomas mioclônicos, ausência, tônico-clônicos, clônicos, tônicos e atônicos.

APÊNDICE 4 – Algoritmos de aprendizagem

I - Aprendizagem com Neurônio Difuso Hirota e Pedrycz (1992, 1994)

Considere um sistema com n unidades de entrada ($u_1, \dots, u_j, \dots, u_n$), q unidades de regras (z_1, z_k, \dots, z_q) e três unidades de saída (y_1, y_2, y_3). Na implementação do algoritmo supervisionado é fornecido, além do conjunto de padrões de entrada (u_j), o conjunto de padrões desejados de saída (d_j) onde $u_j \in \mathbb{R}$ e $d_j \in [0,1]$, de acordo com as seguintes etapas:

• FASE I (*feedforward*):

Etapa_1: selecionar o próximo padrão de *entrada/saída* (u_j, d_j).

Etapa_2: para cada unidade de entrada u_j aplicar a sua função de pertinência (μ) de acordo com seus respectivos CDs. Esta etapa recebe os dados a partir do ambiente externo.

A entrada neural difusa é descrita em termos de graus de pertinência conforme a equação:

$$x_j^p = [\theta_k, x_1, \dots, x_j, \dots, x_n] \in [0,1]$$

onde:

x_j^p = valores de entrada *fuzzificados* do neurônio j para o padrão p ;

θ_k = *bias* da camada intermediária k com os valores +1 na camada intermediária e (-1) na camada de saída¹.

Etapa_3: encontrar as regras apropriadas a partir das unidades *fuzzificadas* de entrada ($x_1, \dots, x_j, \dots, x_n$).

Etapa_4: executar a parte antecedente das regras, por meio do neurônio difuso E . Computar a saída do sinal aplicando a função de ativação logística.

Os valores de x_j e v_{kj} são combinados por uma operação *t-conorma* S para produzir p_k^p :

$$p_k^p = S(v_{kj}^p, x_j^p), j = 1, \dots, n$$

onde:

p_k^p = resultado da operação aritmética difusa *t-conorma* S aplicada no neurônio k da camada intermediária para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p , onde $v_{kj}^p \in [-0.5, 0.5]$;

A entrada da informação p_k^p é agregada por uma operação aritmética difusa *t-norma* T para produzir a saída net_k^p :

$$net_k^p = \theta_k + T(S(p_{11}^p), \dots, S(p_{kj}^p), \dots, S(p_{qn}^p))$$
$$z_k^p = \varphi(net_k^p)$$

¹ O valor negativo do *bias* poderá influenciar no potencial de ativação do neurônio (Haykin, 2001, p. 37).

onde:

net^p_k = entrada da rede atingida pelo neurônio k da camada intermediária para o padrão p ;

$\varphi(net^p_k)$ = função de ativação aplicada na entrada da rede atingida pelo neurônio k da camada intermediária da rede para o padrão p ;

θ^p_k = *bias* da camada intermediária k para o padrão p ;

z^p_k = resultado de saída da camada intermediária k para o padrão p ;

T = operação aritmética difusa *t-norma* (Produto Algébrico, Produto de Hamacher, Produto de Einstein ou a operação aritmética difusa padrão *min*);

S = operação aritmética difusa *t-conorma* (Soma Algébrica, Soma de Hamacher, Soma de Einstein ou a operação aritmética difusa padrão *max*).

Etapa_5: executar a parte antecedente das regras utilizando o neurônio difuso *OU*. Computar a saída do sinal aplicando a função de ativação logística.

Os valores z^p_k e w^p_{ik} são combinados por uma operação *t-norma* T para produzir p^p_i :

$$p^p_i = T(w^p_{ik}, z^p_k), k = 1, \dots, q$$

onde:

w^p_{ik} = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p .

A entrada da informação p^p_i é agregada por uma operação *t-conorma* S para produzir a saída net^p_i :

$$net^p_i = \theta^p_i + S(T(p^p_{1i}), \dots, T(p^p_{qi}), \dots, T(p^p_{mi}))$$

$$y^p_i = \varphi(net^p_i)$$

onde:

θ^p_i = *bias* da camada de saída i para o padrão p ;

$(p^p_{1i}), \dots, (p^p_{qi}), \dots, (p^p_{mi})$ = resultado da operação *t-norma* T aplicada no neurônio i da camada de saída para o padrão p ;

net^p_i = entrada da rede atingida pelo neurônio i da camada de saída para o padrão p ;

$\varphi(net^p_i)$ = função de ativação logística aplicada na entrada da rede atingida pelo neurônio i da camada de saída da rede para o padrão p .

• **FASE II (*backward*):**

Etapa_6: cada unidade de saída recebe a saída desejada, correspondente ao padrão de treinamento, e encontra o seu erro (δ^p_i):

$$\delta^p_i = (d^p_i - y^p_i) * \varphi'(net^p_i)$$

onde:

δ_i^p = erro da unidade de saída i para o padrão p ;

$d_i^p - y_i^p$ = diferença entre a saída desejada (d_i^p) e a saída real (y_i^p) do neurônio i para cada padrão p ;

$\varphi'(net_i^p)$ = derivada da entrada da rede atingida pelo neurônio i da camada de saída para o padrão p .

A seguir são calculados os pesos com o “termo momento”² para depois atualizar w_{ik}^p :

$$\Delta w_{ik}^p = n(\Delta w_{ik}^{p-1}) + \alpha * \delta_i^p * z_k^p$$

onde:

n = constante 1;

Δw_{ik}^{p-1} = variação da intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão $p - 1$;

α = taxa de aprendizagem;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

Após, calcula-se o *bias* utilizado para atualizar θ_i^p :

$$\Delta \theta_i^p = \alpha * \delta_i^p$$

A seguir, envia δ_i^p para as próximas camadas.

onde:

θ_i^p = *bias* da camada de saída i para o padrão p .

Etapa_7: cada unidade intermediária realiza a soma das entradas delta (Δ):

$$\Delta \delta_k^p = \sum \delta_i^p * w_{ik}^p$$

A seguir aplica a derivada de sua função para calcular a formação do erro:

$$\delta_k^p = \Delta \delta_k^p * \varphi'(net_k^p)$$

Realiza o cálculo dos pesos, incluindo o termo momento:

$$\Delta v_{kj}^p = n(\Delta v_{kj}^{p-1}) + \alpha * \delta_k^p * x_j^p$$

E atualiza θ_k^p :

$$\Delta \theta_k^p = \alpha * \delta_k^p$$

Dessa forma, o sinal do erro na saída do neurônio i no padrão p é definido como:

$$e_i^p = (d_i^p - y_i^p) \in [-1, 1]$$

Etapa_8: atualizar as conexões sinápticas (pesos) dos neurônios *OU* da camada de saída³:

$$w_{ik}^{p+1} = w_{ik}^p + \Delta w_{ik}^p$$

onde:

² Variação sugerida por Cichoki (1994), a qual inclui o *termo momento*, tornando a aprendizagem mais eficiente.

³ Os pesos são atualizados de padrão em padrão de acordo com os respectivos erros calculados para cada padrão apresentado à rede (Haykin, 2001).

$$\Delta w_{ik}^p = \alpha * \delta_i^p * z_k^p$$

Etapa_9: atualizar as conexões sinápticas do *bias*:

$$\theta_i^{p+1} = \theta_i^p + \Delta \theta_i^p$$

onde:

$$\Delta \theta_i^p = \alpha * \delta_i^p$$

Etapa_10: atualizar as conexões sinápticas (pesos) dos neurônios E da camada intermediária:

$$v_{kj}^{p+1} = v_{kj}^p + \Delta v_{kj}^p$$

onde:

$$\Delta v_{kj}^p = \alpha * \delta_k^p * x_j^p$$

Etapa_11: atualizar as conexões sinápticas do *bias*:

$$\theta_k^{p+1} = \theta_k^p + \Delta \theta_k^p$$

$$\Delta \theta_k^p = \alpha * \delta_k^p$$

Etapa_12: O próximo passo é calcular o sinal de erro na saída do neurônio *i* para o padrão *p*:

$$e_i^p = (d_i^p - y_i^p)$$

Após definir o erro no padrão *p*, calcular o erro quadrático total *E(t)*:

$$E(t) = (1/2) * \sum (e_i^p)^2$$

Etapa_13: Se existirem mais padrões para o treinamento da rede, voltar à *Etapa_1*.

No final de cada iteração calcular o Erro Médio Quadrático (EMQ) :

$$EMQ(t) = (1/N) * \sum E(t)$$

onde:

N = número de padrões de treinamento.

II - Aprendizagem com Neurônio Difuso Kwan e Cai (1994)

Considere um sistema com n unidades de entrada ($u_1, \dots, u_j, \dots, u_n$), q unidades de regras (z_1, z_2, \dots, z_q) e três unidades de saída (y_1, y_2, y_3). Para a implementação do algoritmo supervisionado é fornecido, além do conjunto de padrões de entrada (u_j), o conjunto de padrões desejados de saída (d_i) onde $u_j \in \mathbb{R}$ e $d_i \in [0,1]$, de acordo com as seguintes etapas:

• FASE I (*feedforward*):

Etapa_1: selecionar o próximo padrão de entrada/saída (u_j, d_i).

Etapa_2: para cada unidade de entrada u_j aplicar a sua função de pertinência (μ) de acordo com seus respectivos CDs. Esta etapa recebe os dados a partir do ambiente externo.

A entrada neural difusa é descrita em termos de graus de pertinência conforme a equação:

$$x_j^p = [\theta_k, x_1, \dots, x_j, \dots, x_n] \in [0,1]$$

onde:

x_j^p = valores de entrada *fuzzificados* do neurônio j para o padrão p ;

θ_k = *bias* da camada intermediária k com os valores +1 na camada intermediária e (-1) na camada de saída.

Etapa_3: encontrar as regras apropriadas de acordo com as unidades *fuzzificadas* de entrada ($x_1, \dots, x_j, \dots, x_n$).

Etapa_4: executar a parte antecedente das regras através do neurônio difuso E . Computar a saída do sinal aplicando a função de ativação logística.

O neurônio difuso E da camada intermediária k apresenta o sinal de entrada x_j interagindo com o peso sináptico v_{kj} para produzir o produto p^p_k :

$$p^p_k = (v_{kj}x_j), j = 1, \dots, n$$

A entrada da informação p^p_k é agregada por uma operação aritmética difusa *t-norma* T para computar a saída net^p_k :

$$net^p_k = \theta^p_k + T\{(p^p_{11}), \dots, (p^p_{kj}), \dots, (p^p_{qn})\} = \theta^p_k + T\{v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n\}$$

onde:

θ^p_k = *bias* da camada intermediária k para o padrão p ;

p^p_k = resultado da operação aritmética produto aplicada no neurônio k da camada intermediária para o padrão p ;

v^p_{kj} = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p ;

x^p_j = valor de entrada *fuzzificado* j para o padrão p .

A saída do neurônio z^p_k é computada por:

$$z_k^p = \varphi(\text{net}_k^p - \theta_k^p) = \varphi(T\{v_{11}x_1, \dots, v_{kj}x_j, \dots, v_{qn}x_n\} - \theta_k^p)$$

onde:

z_k^p = resultado da saída da camada intermediária k para o padrão p ;

$\varphi(\text{net}_k^p)$ = função de ativação aplicada na entrada da rede atingida pelo neurônio k da camada intermediária da rede para o padrão p ;

Etapa_5: executar a parte antecedente das regras utilizando o neurônio difuso *OU*. Computar a saída do sinal aplicando a função de ativação logística.

No neurônio difuso *OU* o sinal de entrada z_k^p interage com o peso sináptico w_{ik} para produzir o produto p_k^p :

$$p_k^p = w_{ik} z_k^p, \quad k = 1, \dots, q$$

A entrada da informação p_k^p é agregada por uma operação *t-conorma* S para produzir a saída net_i^p :

$$\text{net}_i^p = \theta_i^p + S\{(p_{11}^p), \dots, (p_{ik}^p), \dots, (p_{mq}^p)\} = \theta_i^p + S\{w_{11}z_1, \dots, w_{ik}z_k, \dots, w_{mq}z_q\}$$

onde:

θ_i^p = *bias* da camada de saída i para o padrão p ;

p_i^p = resultado da operação aritmética produto aplicada no neurônio i da camada de saída para o padrão p ;

w_{ik}^p = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p ;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

A saída do neurônio y_i^p é computada por:

$$y_i^p = \varphi(\text{net}_i^p - \theta_i^p) = \varphi(S\{w_{11}z_1, \dots, w_{ik}z_k, \dots, w_{mq}z_q\} - \theta_i^p)$$

onde:

y_i^p = camada de saída i para o padrão p .

$\varphi(\text{net}_i^p)$ = função de ativação logística aplicada na entrada da rede atingida pelo neurônio i da camada de saída da rede para o padrão p ;

• **FASE II (backward):**

Etapa_6: cada unidade de saída recebe a saída desejada, correspondente ao padrão de treinamento e encontra o seu erro (δ^p):

$$\delta^p = (d_i^p - y_i^p) * \varphi'(\text{net}_i^p)$$

onde:

δ^p = erro da unidade de saída i para o padrão p ;

$d^p_i - y^p_i$ = diferença entre a saída desejada (d^p_i) e a saída real (y^p_i) do neurônio i para cada padrão p .

A seguir são calculados os pesos utilizando o “termo momento” para depois atualizar w^p_{ik} :

$$\Delta w^p_{ik} = n(\Delta w^{p-1}_{ik}) + \alpha * \delta^p_i * z^p_k$$

onde:

n = constante 1;

w^{p-1}_{ik} = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão $p - 1$;

α = taxa de aprendizagem;

z^p_k = resultado de saída da camada intermediária k para o padrão p .

Após, calcula-se o *bias* utilizado para atualizar θ^p_i :

$$\Delta \theta^p_i = \alpha * \delta^p_i$$

A seguir, envia δ^p_i para as próximas camadas.

onde:

θ^p_i = *bias* da camada de saída i para o padrão p .

Etapa_7: a descrição dessa etapa até a *Etapa_13* segue de acordo com a descrição do neurônio difuso proposto por Hirota e Pedrycz.

III - Aprendizagem para as Funções *max* e *min* (Nauck, 1994)

Considere-se um sistema com n unidades de entrada ($u_1, \dots, u_j, \dots, u_n$), q unidades de regras (z_1, z_2, \dots, z_q) e três unidades de saída (y_1, y_2, y_3). Para a implementação do algoritmo supervisionado é fornecido, além do conjunto de padrões de entrada (u_j), o conjunto de padrões desejados de saída (d_j), onde $u_j \in \mathbf{R}$ e $d_j \in [0,1]$.

O algoritmo de aprendizagem é implementado para criar q unidades de regras, de acordo com as etapas adaptadas de NAUCK (1994):

· FASE I (*feedforward*):

*Etapas*_1: selecionar o padrão de *entrada/saída* (u_j, d_j);

*Etapas*_2: para cada unidade de entrada u_j aplicar a sua função de pertinência (μ) de acordo com seus respectivos CDs. Esta etapa recebe os dados a partir do ambiente externo.

A entrada neural difusa é descrita em termos de graus de pertinência conforme a equação:

$$x_j^p = [\theta_k, x_1, \dots, x_j, \dots, x_n] \in [0,1]$$

onde:

x_j^p = valores de entrada *fuzzificados* do neurônio j para o padrão p ;

θ_k = *bias* da camada intermediária k com os valores +1 na camada intermediária e (-1) na camada de saída.

*Etapas*_3: encontrar as regras apropriadas de acordo com as unidades *fuzzificadas* de entrada ($x_1, \dots, x_j, \dots, x_n$).

*Etapas*_4: executar a parte antecedente das regras utilizando a operação aritmética difusa padrão *min* aplicando o neurônio difuso E . Os valores x_j^p e v_{kj}^p são combinados por uma operação *t-norma (min)* para produzir p_k^p :

$$p_k^p = \min(v_{kj}^p, x_j^p)$$

onde:

p_k^p = resultado da operação *t-norma (min)* aplicada no neurônio k da camada intermediária para o padrão p ;

v_{kj}^p = intensidade da conexão do neurônio j da camada de entrada para o neurônio k da camada intermediária no padrão p , onde $v_{kj}^p \in [-0.5, 0.5]$;

À entrada da informação p_k^p são acrescentados o *bias* e seu respectivo peso para produzir a saída net^p_k :

$$net^p_k = \theta_k + (w_{k1}^p), \dots, (w_{kj}^p), \dots, (w_{kn}^p)$$

$$z_k^p = \varphi(net^p_k)$$

onde:

net^p_k = entrada da rede atingida pelo neurônio k da camada intermediária para o padrão p ;

$\varphi(net^p_k)$ = função de ativação aplicada na entrada da rede atingida pelo neurônio k da camada intermediária da rede para o padrão p ;

θ^p_k = *bias* da camada intermediária k para o padrão p ;

z^p_k = resultado de saída da camada intermediária k para o padrão p .

Etapa_5: executar a parte antecedente das regras utilizando a operação aritmética difusa padrão *max* aplicando o neurônio difuso *OU*. Os valores z^p_k e w^p_{ik} são combinados por uma operação *t-conorma (max)* para produzir p^p_i :

$$p^p_i = \max (w^p_{ik}, z^p_k)$$

onde:

p^p_i = resultado da operação *t-conorma (max)* aplicada no neurônio i da camada de saída para o padrão p ;

w^p_{ik} = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão p .

À entrada da informação p^p_i são acrescentados o *bias* e seu respectivo peso para produzir a saída net^p_i :

$$net^p_i = \theta^p_i + ((p^p_{1k}), \dots, (p^p_{ik}), \dots, (p^p_{mk}))$$
$$y^p_i = \varphi(net^p_i)$$

onde:

θ^p_i = *bias* da camada de saída i para o padrão p ;

$(p^p_{1k}), \dots, (p^p_{ik}), \dots, (p^p_{mk})$ = resultado da operação aritmética difusa *t-norma T* aplicada no neurônio i da camada de saída para o padrão p ;

net^p_i = entrada da rede atingida pelo neurônio i da camada de saída da rede para o padrão p ;

$\varphi(net^p_i)$ = função de ativação logística aplicada na entrada da rede atingida pelo neurônio i da camada de saída da rede para o padrão p .

• **FASE II (*backward*):**

Etapa_6: cada unidade de saída recebe a saída desejada, correspondente ao padrão de treinamento e encontra o seu erro δ^p_i :

$$E^p = 1 - \exp \left(-\beta \left(\frac{(d^p_i - y^p_i)}{y^p_i(\text{maximo}) - y^p_i(\text{minimo})} \right)^2 \right)$$

$$\delta_i^p = \text{sgn}(d_i^p - y_i^p) * E^p$$

onde:

E^p = erro no padrão p ;

sgn = função sinal;

$d_i^p - y_i^p$ = diferença entre a saída desejada (d_i^p) e a saída real (y_i^p) do neurônio i para cada padrão p ;

$y_i^p(\text{maximo}) - y_i^p(\text{minimo})$ = diferença entre o valor de saída real máximo e o valor de saída real mínimo (y_i^p) do neurônio i para cada padrão p ;

β = é um fator escalar;

δ_i^p = erro da unidade de saída i para o padrão p .

A seguir são calculados os pesos utilizando o “termo momento” para depois atualizar w_{ik}^p :

$$\Delta w_{ik}^p = n(\Delta w_{ik}^{p-1}) + \alpha * \delta_i^p * z_k^p$$

onde:

n = constante 1;

w_{ik}^{p-1} = intensidade da conexão do neurônio k da camada intermediária para o neurônio i da camada de saída no padrão $p - 1$;

α = taxa de aprendizagem;

z_k^p = resultado de saída da camada intermediária k para o padrão p .

Após, calcula-se o *bias* utilizado para atualizar θ_i^p :

$$\Delta \theta_i^p = \alpha * \delta_i^p$$

A seguir, envia δ_i^p para as próximas camadas.

onde:

θ_i^p = *bias* da camada de saída i para o padrão p .

Etapa_7: cada unidade intermediária realiza a soma das entradas delta (Δ) para encontrar a formação do erro:

$$\Delta \delta_k^p = \sum \delta_i^p * w_{ik}^p$$

Realiza o cálculo dos pesos, incluindo o termo momento:

$$\Delta v_{kj}^p = n(\Delta v_{kj}^{p-1}) + \alpha * \delta_k^p * x_j^p$$

E atualiza θ_k^p :

$$\Delta \theta_k^p = \alpha * \delta_k^p$$

Dessa forma, o sinal do erro na saída do neurônio i no padrão p é definido como:

$$e_i^p = (d_i^p - y_i^p) \in [-1, 1]$$

Etapa_8: a descrição dessa etapa até a Etapa_13 segue de acordo com a descrição do neurônio difuso proposto por Hirota e Pedrycz.

APÊNDICE 5 – Grafos do Conhecimento

I – Eventos Epilépticos (Ilae, 1981)

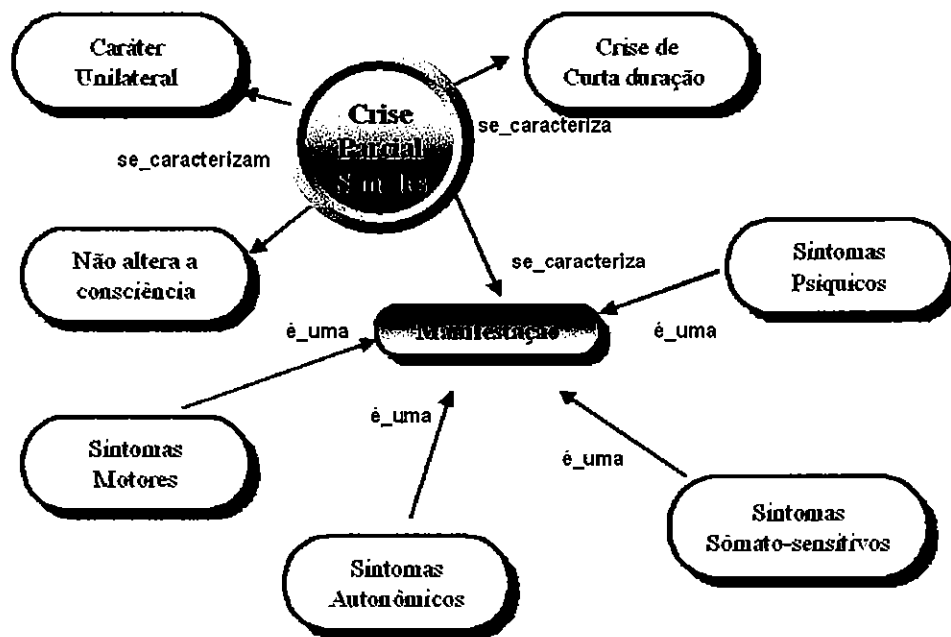


Figura 1 – Crise parcial simples

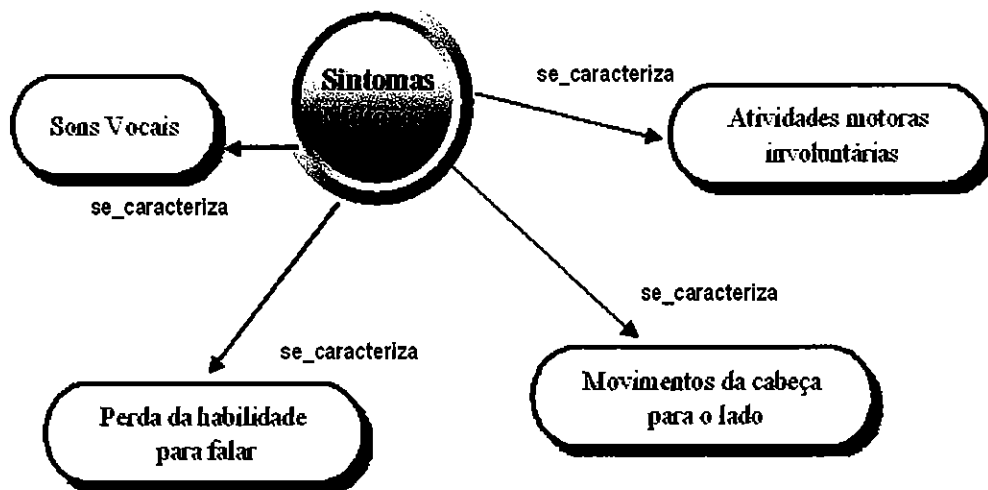


Figura 2 – Sintomas motores

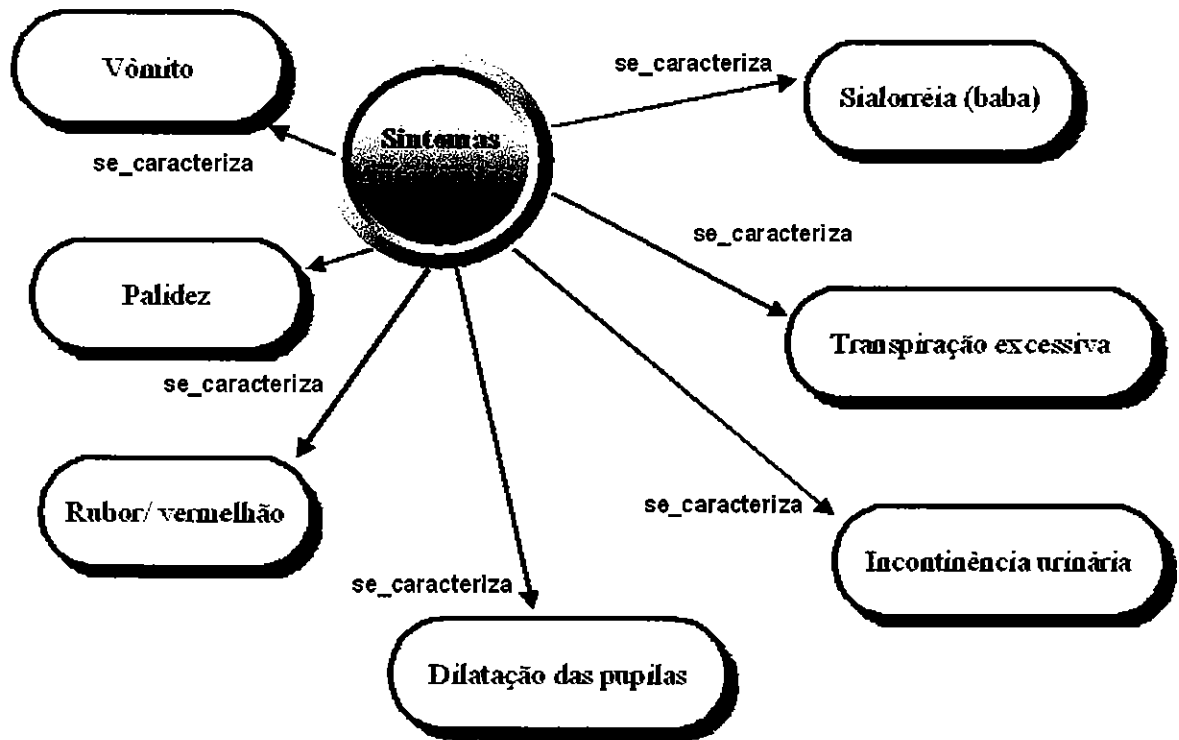


Figura 3 – Sintomas autonômicos

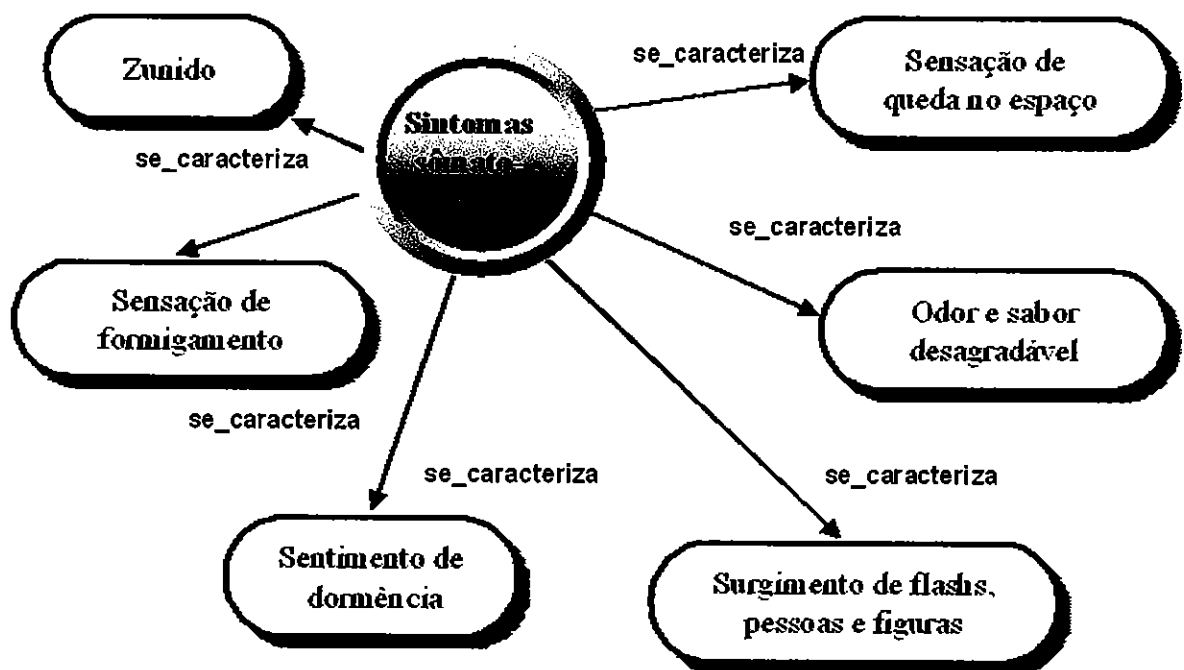


Figura 4 – Sintomas sômato-sensitivos

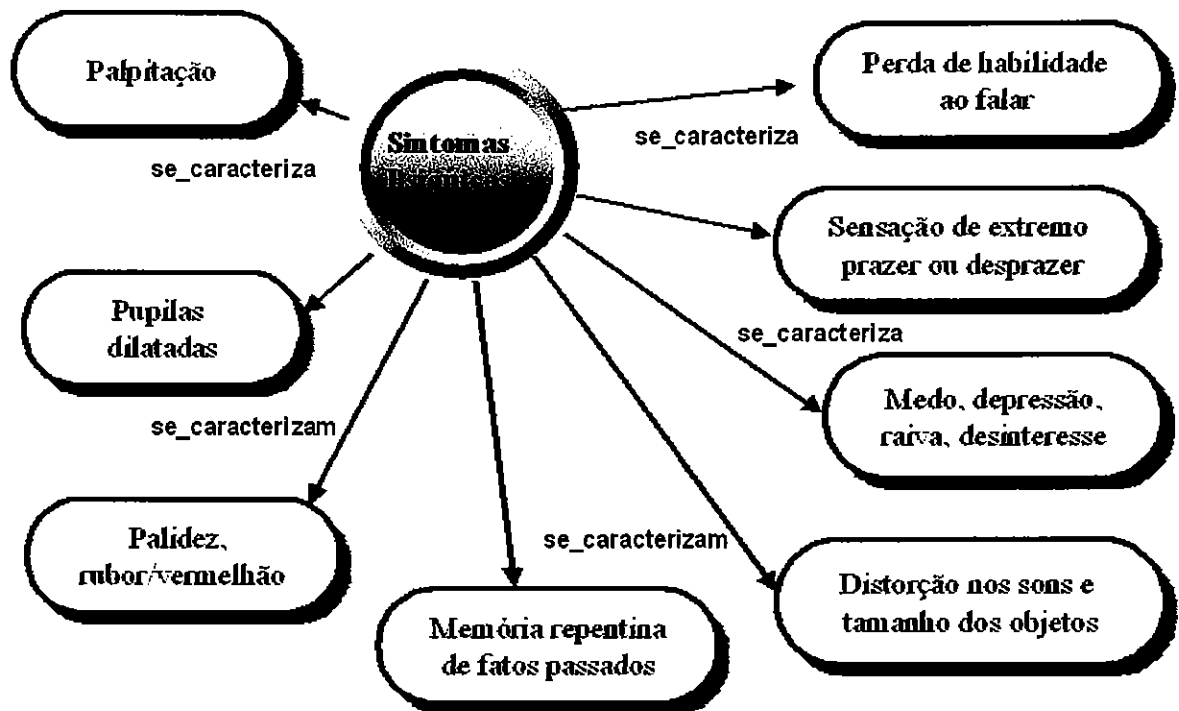


Figura 5 – Sintomas psíquicos

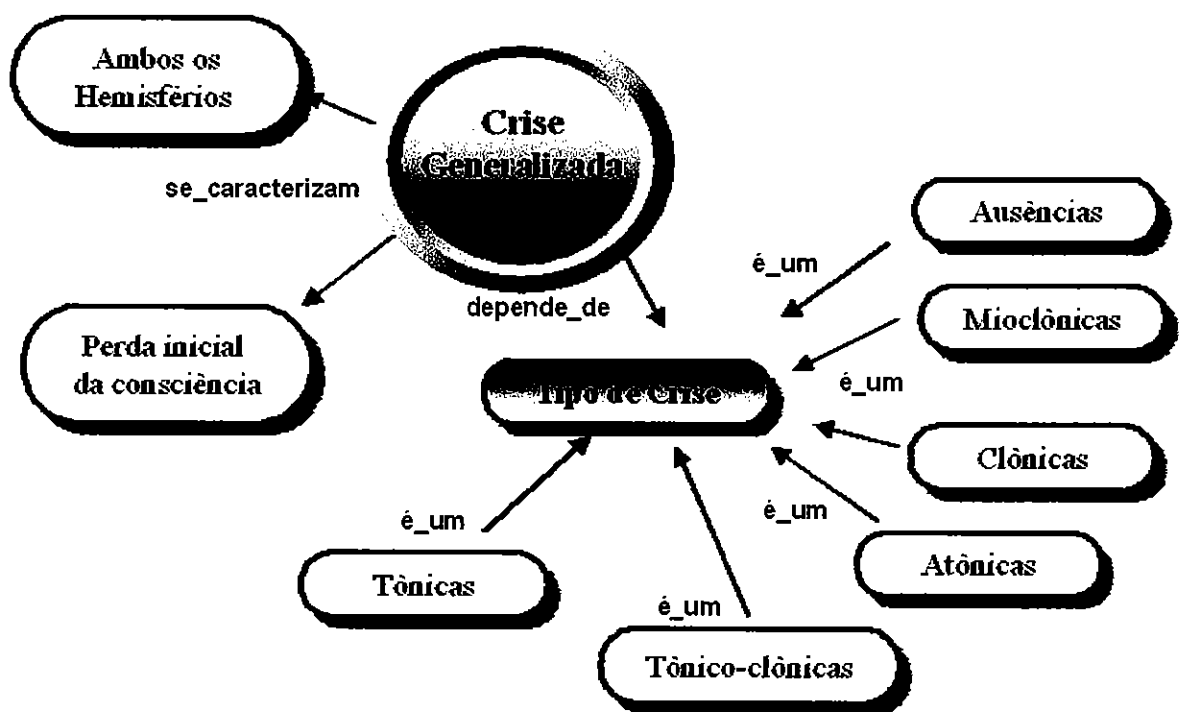


Figura 6 – Crise generalizada

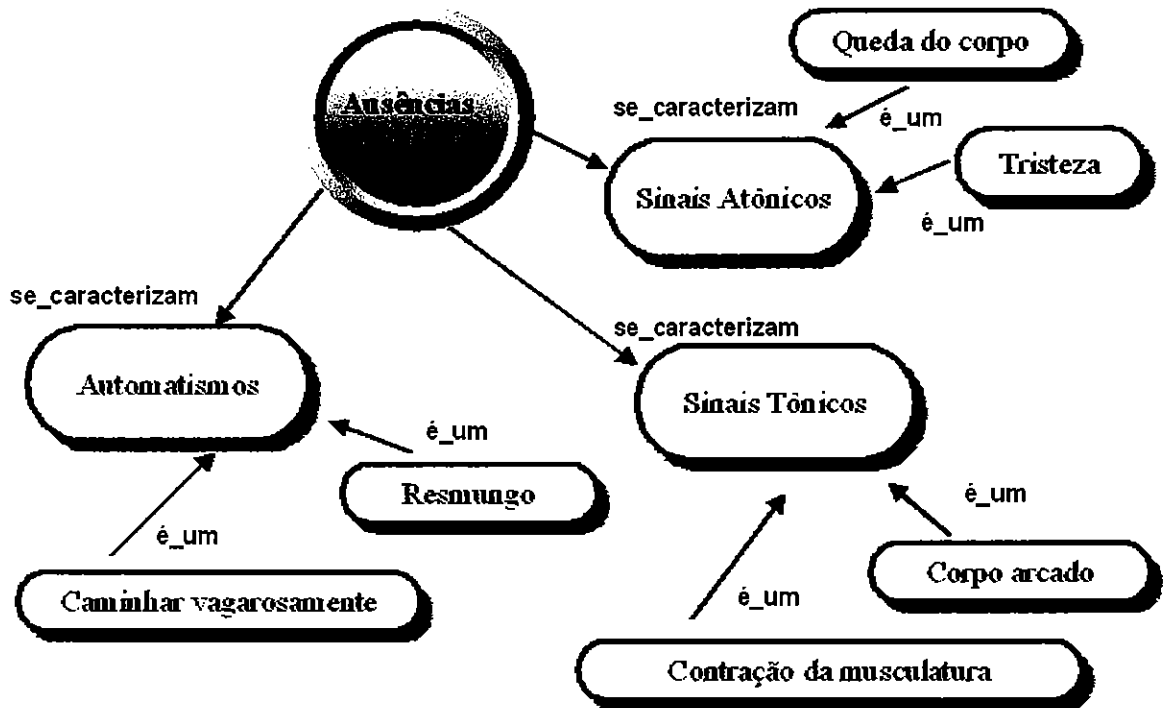


Figura 7– Ausências atípicas

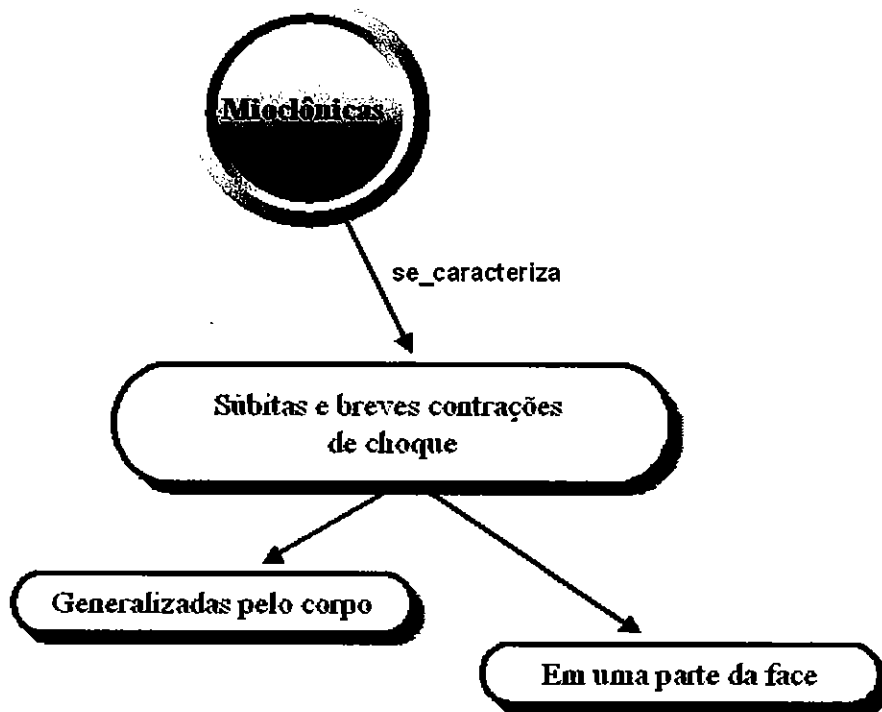


Figura 8– Crises mioclônicas

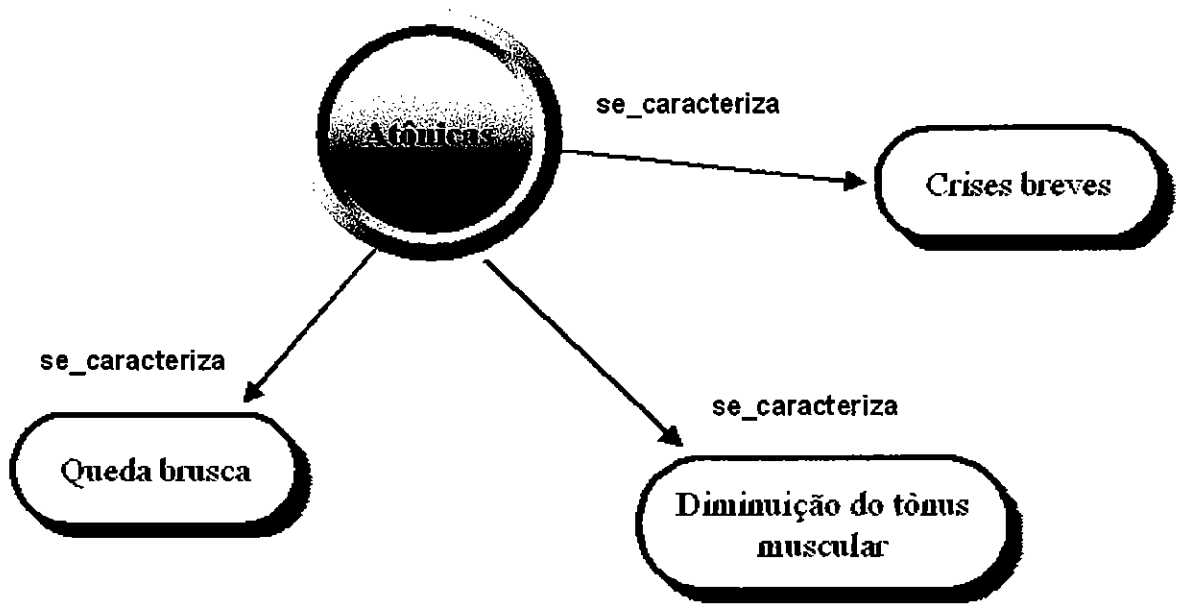


Figura 9 – Crises atônicas

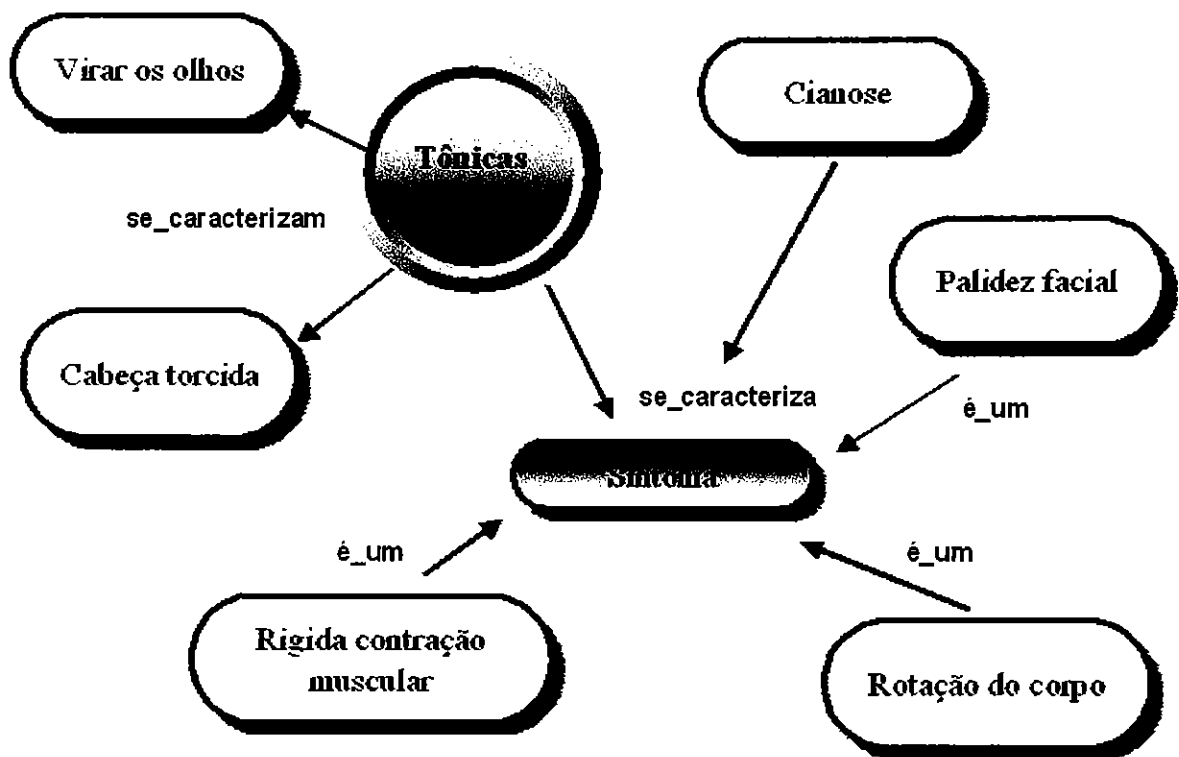


Figura 10 – Crises tônicas

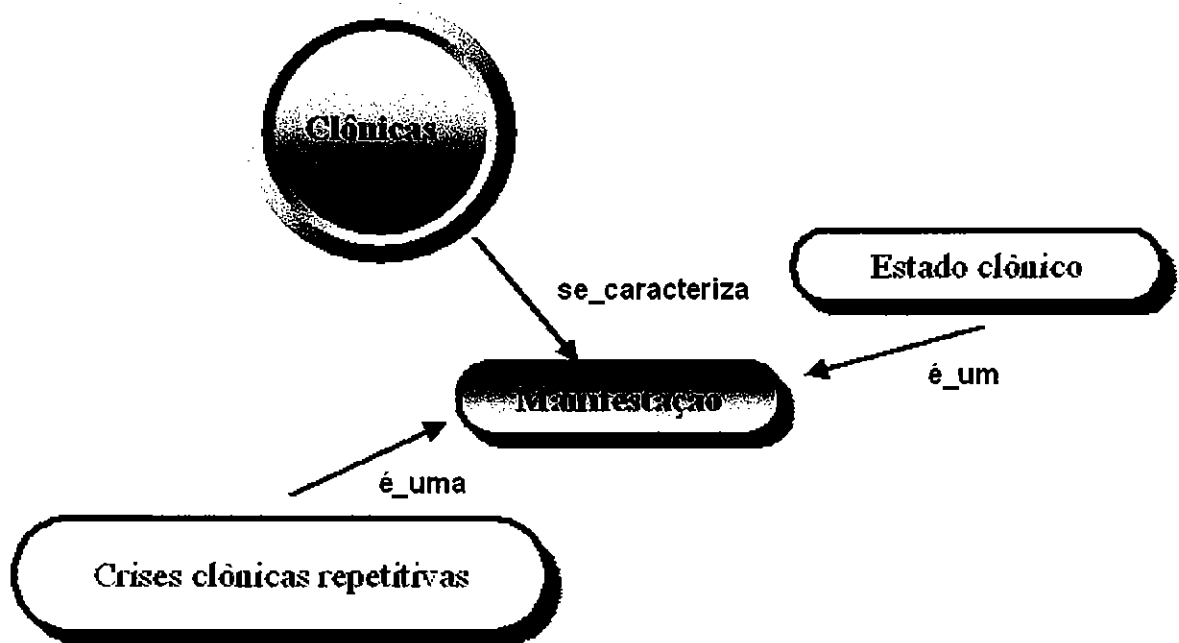


Figura 11 – Crises clônicas

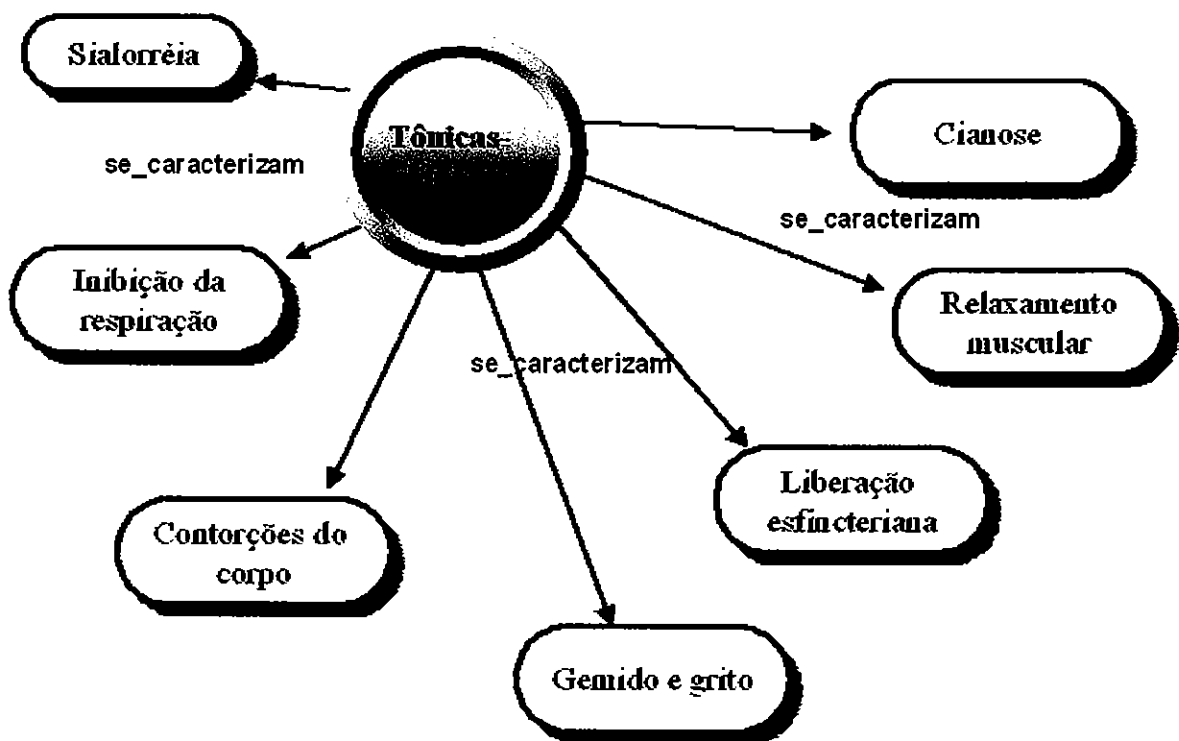


Figura 12 – Crises tônico-clônicas

II – Eventos Não Epilépticos (Costa et al., 1998)

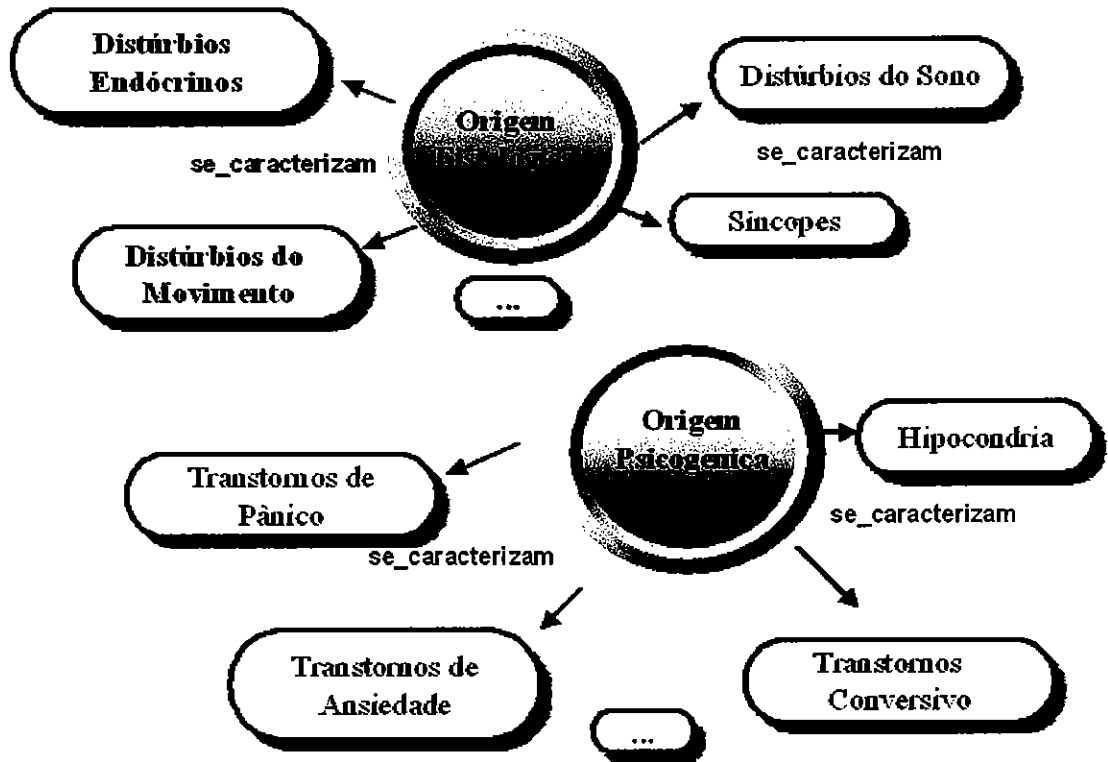


Figura 13 - Eventos não epilépticos de origem fisiológica e psicogênica

III – Eventos Não Epilépticos (Luther et al., 1982)

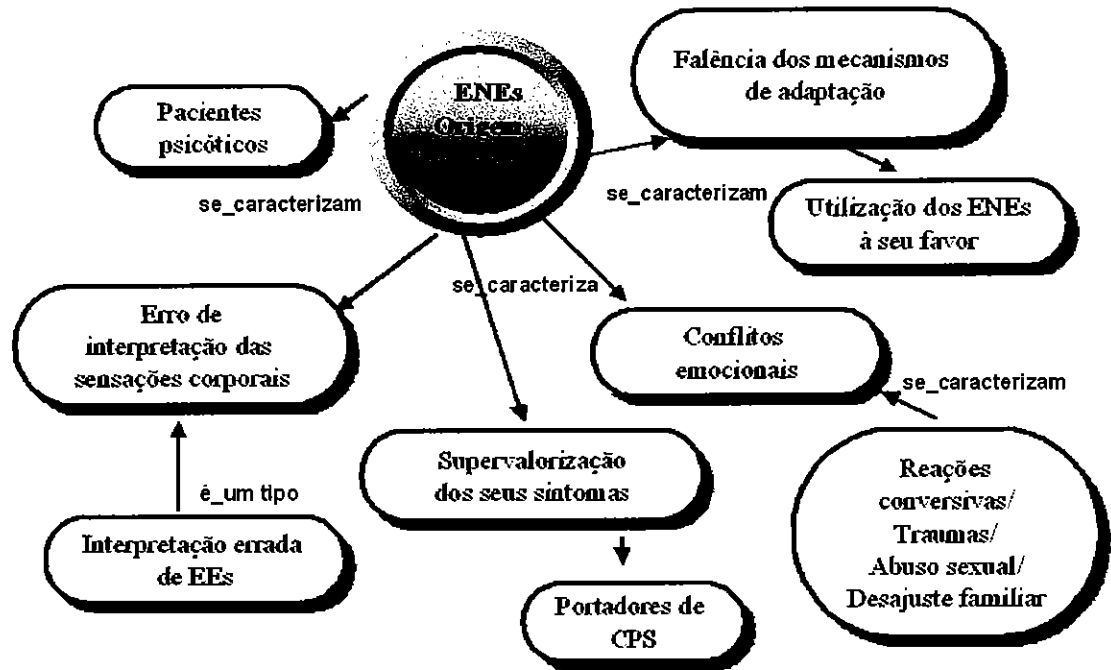


Figura 14 – ENEs de origem psicogênica

IV – Eventos Não Epilépticos (Freud, 1998)

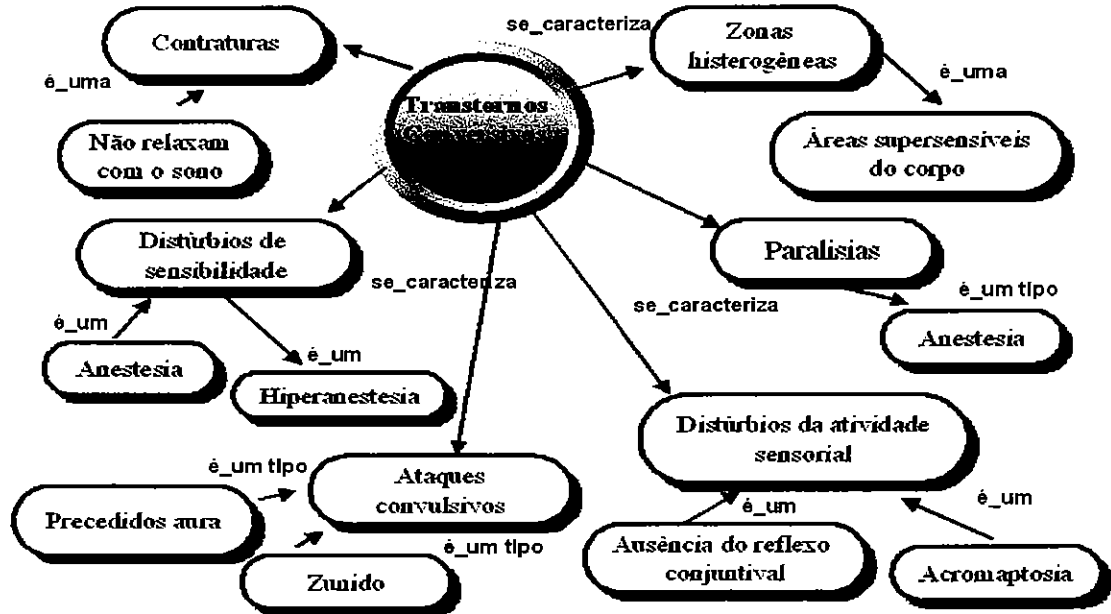


Figura 15 – Transtornos conversivos

APÊNDICE 6 - Glossário

Estas definições estão baseadas em Abc da Saúde (2005), Cecil (1990), Costa et al. (1998), Fernandes (1993), Ilae (1981), Larouse (1999).

- Amaurose – perda total da visão.
- Aura – do grego, significa *brisa*. Tem sido usado para designar o início de uma crise epiléptica. Toda a aura, de qualquer tipo, pode ser considerada uma crise parcial simples.
- Automatismos – atos comportamentais involuntários que ocorrem na presença de alterações da consciência, durante ou após uma crise.
- Centelhas – partícula ígnea ou flashes luminosos que saem de um corpo incandescente.
- Cianose – ficar com a pele com coloração azul.
- Compressão na garganta – aperto, constrição na garganta.
- Contraturas - ato ou efeito de contrair-se. Não relaxam com o sono, como ocorre com as contraturas orgânicas e não se consegue modificar sua intensidade mediante excitação, temperatura, etc. Somente recobram sua intensidade total quando o paciente acorda.
- Crise de ausência - surgimento de automatismos de atenção, mímica com gestos de arrumar a roupa e olhar fixo no espaço vazio. Breves interrupções, entre elas: rotação dos olhos para cima, voz, caminhada, alimentares e interrupções da escuta;
- Crise de ausência atípica - com suaves movimentos clônicos da pálpebra e canto da boca. Os objetos são derrubados. Paciente poderá apresentar: (1) sinais atônicos: diminuição do tônus muscular, tristeza, queda do corpo, caminhar de braços abertos; (2) sinais tônicos: contração da musculatura, contorcimento da cabeça para trás ou para os lados, corpo arcado e (3) automatismos: caminhar vagorosamente, ranger os lábios e resmungo.
- Crise atônica - caracterizadas por crises breves, diminuição do tônus muscular, queda do corpo ou cabeça para um lado e queda brusca.
- Crise clônica - caracterizadas por crises clônicas repetitivas, ou seja, o paciente deixa-se cair em estado clônico.
- Crise mioclônica - súbitas e breves contrações de choque generalizadas pelo corpo ou numa parte da face, podendo ocorrer durante o sono ou em vigília, repetidamente ou isoladas.

- Crise parcial complexa - atividades motoras involuntárias seguidas de eventos de anamnesia, apresentando os seguintes automatismos: alimentares, arremedação/mímica, gesticulares e verbais. O paciente apresenta a tendência de arrumar várias vezes as suas próprias roupas.
- Crise tônica - rígida contração muscular, virar os olhos, cabeça torcida para trás ou para os lados, rotação do corpo, palidez facial e cianose.
- Crise tônico-clônica - súbita contração tônica dos músculos, aparecimento de som agudo/áspero, gemido, grito, deixando o paciente cair em estado clônico. Contorções do corpo para trás, inibição da respiração, cianose (ficar roxa), sabor amargo na língua, liberação esfinteriana (paciente poderá se urinar), grunidos respiratórios, sialorréia (baba), respiração profunda e relaxamento muscular. Após a crise, o paciente sente dores no corpo, dor de cabeça e poderá cair em sono profundo.
- Despersonalização - mudar a personalidade, o caráter.
- Déjà vu/jamais vu - sensação de já ter visto e/ou não alguma coisa.
- Disfasia – incapacidade para ordenar as palavras.
- Dissociativo – ação de dissociar-se, perversão de costumes, devasso.
- Distúrbios da atividade sensorial: estes distúrbios podem afetar todos os órgãos dos sentidos. Seus sintomas característicos são: fundo de olho normal ao exame, ausência do reflexo conjuntival (reflexo corneano diminuído); estreitamento concêntrico do campo visual; redução da percepção luminosa; e acromaptosia, isto é o paciente enxerga na cor preta, cinza e branca. No caso do último sintoma citado, a sensibilidade ao roxo é a primeira a ser perdida, e a sensibilidade ao vermelho ou ao azul, é a que persiste por mais tempo. Os objetos, que se aproximam do olho e que dele se afastam, são vistos em tamanhos diferentes, duplicados ou multiplicados (macropsia ou micropsia). A surdez histérica raramente é bilateral; é, com muita frequência, mais ou menos completa, combinada com anestesia do pavilhão da orelha, do canal auditivo e até mesmo da membrana do tímpano;
- Distúrbios da sensibilidade – anestesia; podem afetar a pele, as membranas mucosas, os ossos, os músculos e nervos, os órgãos dos sentidos e os intestinos. A anestesia pode ser total ou atingir apenas a sensação de dor (analgesia - que é a mais comum).
- Distúrbios endócrinos – das glândulas de secreção interna, ou referente a elas.
- Espasmos da faringe – contração dos músculos da faringe, como se uma bola estivesse subindo do epigástrico para a garganta.

- Flash - luminosidade intensa e instantânea.
- Hemianopsia - perda da metade do campo visual.
- Hipocondria – estado mental em que há depressão e doentia preocupação com a própria saúde.
- Parestesias – sensação tátil anormal, como por exemplo, queimação ou de formigamento.
- Paroxística - início e término abrupto.
- Prodrômicos – sinais que precedem ocasionalmente as crises, podendo incluir uma alteração do humor, sensação de apreensão, insônia ou perda do apetite.
- Sialorréia – aumento anormal do fluxo salivar; hipersalivação.
- Síncope – perda breve e repentina da consciência, ocorre quando a circulação cerebral diminui. O paciente costuma sentir náuseas, tonturas, suor, palidez, visão borrada/acinzentada. A síncope ocorre quase sempre na posição sentada e não causam nenhuma confusão mental ou cefaléia pós-crise.
- Sintomas motores: atividades motoras involuntárias, movimentos da cabeça para um lado apenas, perda da habilidade de falar, surgimento de sons vocais;
- Sintomas autonômicos: vômito, palidez, rubor/vermelhão, dilatação das pupilas, incontinência urinária, transpiração excessiva, sialorréia (baba);
- Sintomas somatossensoriais: sensação de formigamento (parestesias), sentimento de dormência, surgimento de flachs, pessoas e figuras, zunido, odor e sabor desagradável, sensação de queda no espaço;
- Sintomas físicos: perda da habilidade de falar, memória repentina de um fato passado, falar que conhece alguém ou lugar sem nunca ter conhecido ou estado lá, falar que jamais compreendeu alguma coisa ou que compreende algo, estados de sonhos, desinteresse, terror, sensação de extremo prazer ou desprazer, medo, intensa depressão, raiva ou fúria, pupilas dilatadas, palidez, rubor/vermelhão, palpitação, ataques de risada, distorção no tamanho dos objetos (macropsia/micropsia), distorção de sons, surgimento de ruídos rápidos, despersonalização das pessoas (alteração no tamanho de partes do corpo).
- Somatização – converter experiência ou estado mental em sintoma orgânico.
- Sudorese – secreção de suor.
- Transtornos de pânico – susto ou pavor repentino.

- Transtorno de somatização – converter uma experiência ou estado mental em sintoma orgânico.

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)