

UNIVERSIDADE DE BRASÍLIA – UNB
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE PÓS-GRADUAÇÃO EM ENGENHARIA
MECÂNICA

CARACTERIZAÇÃO E DETECÇÃO AUTOMÁTICA DE
EVENTOS EPILEPTIFORMES EM SINAIS DE
ELETROENCEFALOGRAMA POR DINÂMICA
SIMBÓLICA

LYNWOOD LIVI DE SOUZA

ORIENTADOR: TITO DIAS JUNIOR

DISSERTAÇÃO DE MESTRADO EM CIÊNCIAS MECÂNICAS

BRASÍLIA/DF: OUTUBRO – 2012

UNIVERSIDADE DE BRASÍLIA – UNB
FACULDADE DE TECNOLOGIA
DEPARTAMENTO PÓS-GRADUAÇÃO EM ENGENHARIA
MECÂNICA

CARACTERIZAÇÃO E DETECÇÃO AUTOMÁTICA DE EVENTOS
EPILEPTIFORMES EM SINAIS DE ELETROENCEFALOGRAMA
POR DINÂMICA SIMBÓLICA
LYNWOOD LIVI DE SOUZA

DISSERTAÇÃO SUBMETIDA AO DEPARTAMENTO DE
ENGENHARIA MECÂNICA DA FACULDADE DE TECNOLOGIA
DA UNIVERSIDADE DE BRASÍLIA COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE
MESTRE EM CIÊNCIAS MECÂNICAS.

APROVADA POR:

Prof. Tito Dias Junior (ENM-UnB)
(Orientador)

Prof. Mario Olavo Magno de Carvalho (ENM-UnB)
(Examinador Interno)

Profa. Suélia de Siqueira Rodrigues Fleury Rosa
(Examinadora Externa)

BRASÍLIA/DF, 05 DE OUTUBRO DE 2012

FICHA CATALOGRÁFICA

SOUZA, LYNWOOD LIVI DE

Caracterização e detecção automática de eventos epileptiformes em sinais de eletroencefalograma por dinâmica simbólica [Distrito Federal] 2012.

81 p., 210 x 297 mm (ENM/FT/UnB, Mestre, Ciências Mecânicas, 2012).
Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Mecânica.

1. Eletroencefalograma

2. Dinâmica Simbólica

3. Redes Neurais

4. Algoritmos Genéticos

I. ENM/FT/UnB

II. Título (MESTRE)

REFERÊNCIA BIBLIOGRÁFICA

SOUZA., L.L. (2012). Caracterização e detecção automática de eventos epileptiformes em sinais de eletroencefalografia por dinâmica simbólica. Dissertação de Mestrado em Ciências Mecânicas, Publicação ENM.DM-179A/2012, Departamento de Engenharia Mecânica, Universidade de Brasília, Brasília, DF, 81 p.

CESSÃO DE DIREITOS

AUTOR: Lynwood Livi de Souza.

TÍTULO: Caracterização e detecção automática de eventos epileptiformes em sinais de eletroencefalografia por dinâmica simbólica.

GRAU: Mestre

ANO: 2012

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Lynwood Livi de Souza

Avenida Coronel Gaspar nº 117, Centro.

76.400-000 Uruaçu – GO – Brasil.

DEDICATÓRIA

Dedico esse trabalho a memória de meu amado pai, o qual foi importantíssimo em minha vida e trajetória para conclusão deste trabalho, oferecendo sempre seu apoio, incentivo e presença, deixando enfim, uma saudade infinita.

AGRADECIMENTOS

Ao meu orientador Tito Dias Junior por ser essa pessoa tão especial e ter proporcionado a infinita honra de estar integrando esse trabalho.

À minha mãe, irmã e sobrinhos por serem tão especiais e trazerem a certeza de lutar para uma vida cada vez melhor.

À minha noiva Thiane Marques, por todo carinho e apoio em toda esta trajetória.

Aos professores e funcionários do Departamento de Engenharia Mecânica da Universidade de Brasília pelo apoio e formação acadêmica.

Ao meu grande amigo Gilderlan Girão Pereira, pela sua generosidade, receptividade e simpatia.

À Isabel Rodrigues, Assistente Administrativa do programa, por todo apoio e empenho a mim e ao meu orientador para que se tornasse possível a defesa desse trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo auxílio financeiro ao longo deste trabalho.

RESUMO

O presente trabalho apresenta uma nova metodologia para caracterização, detecção e classificação de sinais de exames de eletroencefalografia (EEG), visando o aprimoramento e agilidade na obtenção de diagnósticos com o objetivo de auxiliar o profissional de saúde, uma vez que os exames de EEG são caracterizados por longos trechos de registros temporais da atividade elétrica do cérebro, que devem ser minuciosamente examinados pelos médicos. Devido à dificuldade associada à caracterização e à detecção de sinais epileptiformes e à importância do diagnóstico, existem na literatura vários métodos desenvolvidos para abordar este problema, tais como: Métodos Auto-Regressivos (AR), Aproximação da Energia (TEO), Análise de Domínio de Frequência, Análise de Tempo-Frequência, Transformada de Wavelets e Métodos Não-Lineares como os modelos estatísticos. Porém, nenhum modelo não foi capaz de atingir uma performance inteiramente confiável em seus resultados, onde a média de acertos entre os trabalhos expostos na literatura fica em torno de 80,7%, um índice de acertos que apesar de significativo, ainda não é satisfatório para este tipo de exame. A fim de contribuir para o estudo do problema, propôs-se neste trabalho a aplicação da Dinâmica Simbólica para caracterização dos sinais, que se baseia na representação por seqüências de símbolos do estado do sistema e operadores de mudança de estado, e algoritmos genéticos para a otimização da representação, de forma que os sinais epileptiformes pudessem ser distinguidos dos sinais considerados normais por uma rede neural artificial, treinada para este fim. Os resultados obtidos demonstram que a metodologia fornece 92,4% de precisão e 96% de acerto.

ABSTRACT

This paper presents a new technique for characterization and classification of signals of tests electroencephalography (EEG), seeking to improve agility and make a diagnosis in order to assist the professional, since the EEG tests are characterized by long stretches of time records electrical activity of the brain that should be thoroughly examined by doctors. Due to the difficulty associated with the characterization and detection of epileptiform signs and the importance of diagnosis exist in literature several methods developed to address this problem, such as autoregressive method (AR), Energy Approach (TEO), Domain Analysis Frequency, Time-Frequency Analysis, Wavelet Transform and Non-Linear methods such as statistical models. But no model has not been able to achieve a performance fully confident in their results, where the mean score among the works exhibited in the literature is around 80.7%, although a number is not a significant acceptable value for this type of examination. To contribute to the study of the problem, proposed in this paper the application of symbolic dynamics for the characterization of signals, which is based on representation by sequences of symbols of the state of the system and operators of state change, and genetic algorithms for optimization of the representation, so that the signals could be distinguished from epileptiform signals considered normal for an artificial neural network trained for this purpose. The results demonstrate that the method provides 92.4% precision and 96% accuracy.

SUMÁRIO

1.	INTRODUÇÃO.....	1
2.	OBJETIVO E METAS	4
2.1.	OBJETIVO	4
2.2.	METAS.....	4
3.	ORGANIZAÇÃO.....	5
4.	REVISÃO BIBLIOGRAFICA	6
4.1.	DINÂMICA E ANÁLISE DE SINAIS BIOELÉTRICOS.....	6
4.2.	AQUISIÇÃO DOS SINAIS BIOELÉTRICOS.....	8
4.3.	INTRODUÇÃO AO SINAL DE EEG	9
4.3.1.	Morfologia do Sinal EEG	11
4.4.	DINÂMICA SIMBÓLICA	15
4.4.1.	Introdução.....	15
4.4.2.	Dinâmica de Sistemas.....	20
4.5.	ANÁLISE DE SINAIS CAÓTICOS	21
4.5.1.	Introdução.....	21
4.6.	REDES NEURAIIS ARTIFICIAIS PARA CLASSIFICAÇÃO.....	22
4.7.	ALGORITMOS GENÉTICOS	30
4.7.1.	Introdução.....	30
4.7.2.	Fundamentos Biológicos	31
4.7.3.	Características globais dos algoritmos genéticos	33
4.7.4.	Operadores Genéticos.....	36

4.7.5. Parâmetros Genéticos	37
4.7.6. Pesquisas e Aplicações sobre AG.....	38
5. METODOLOGIA.....	40
5.1. OBTENÇÃO DOS DADOS DE EEEG	40
5.2. SIMBOLIZAÇÃO DAS SÉRIES.....	40
5.3. CLASSIFICAÇÃO DOS SINAIS.....	44
5.4. ÍNDICES DE DESEMPENHO.....	48
5.5. MODELAGEM DO SOFTWARE	50
6. RESULTADOS	51
6.1. APRESENTAÇÃO DO SOFTWARE.....	51
6.2. SIMBOLIZAÇÃO DE SINAIS PERIÓDICOS E CAÓTICOS	57
6.3. INFLUÊNCIA DO RUÍDO NA SIMBOLIZAÇÃO.....	62
6.4. SIMBOLIZAÇÃO DE EEG NORMAIS E EPILÉPTICOS	63
6.5. TREINAMENTO DA REDE NEURAL PARA A CLASSIFICAÇÃO.....	68
7. CONCLUSÕES E DISCUSSÕES	72
8. SUGESTÕES PARA TRABALHOS FUTUROS.....	75
9. REFERÊNCIAS BIBLIOGRÁFICAS	76

LISTA DE TABELAS

Tabela 4.1 - Fonte: Webster J. G., Editor: Medical Instrumentation, Application and Designer. Wiley, 1998.	8
Tabela 4.2– Ritmo das principais ondas cerebrais (Extraído de PARREIRA, 2006).	14
Tabela 4.3 – Exemplo aleatório da frequência das sequências.....	16
Tabela 5.2 – Tabela da usabilidade dos algoritmos para detecção de crises (Extraído de PARREIRA, 2006).	44
Tabela 5.3 – Redes neurais mais utilizadas – (Extraído de PARREIRA, 2006).	44
Tabela 5.4 – Tabela dos perfis dos sinais.	48
Tabela 6.1 – Valores dos operadores genéticos.....	58
Tabela 6.2 – Valores otimizados pelo Algoritmo Genético.	60
Tabela 6.5 – Valores dos parâmetros otimizados na simbolização das séries temporais utilizando o Algoritmo genético.	66
Tabela 6.6 – Parâmetros para construção da rede MLP.	68
Tabela 6.7 – Conjunto de treinamento para rede MLP.	69
Tabela 6.8 – Resultados dos teste obtidos neste trabalho.....	70
Tabela 6.9 – Informações dos arquivos utilizados para compor da base de dados	71

LISTA DE FIGURAS

Figura 1.1 – Registro de atividade de eletroencefalográfica. Fonte: http://www.cerebromente.org.br/n03/tecnologia/eeg_p.htm	2
Figura 4.1 – Sinais bioelétricos retirados de diferentes partes do corpo: a) ECG, b) EEG, c) EMG e d) EOG. (Fonte: IEM0700 Biomedical Electronics, <a href="http://www.elin.ttu.ee/mesel/Study/Courses/Biomedel/Content/BioElect/BE
Signal/BEsignal.htm">http://www.elin.ttu.ee/mesel/Study/Courses/Biomedel/Content/BioElect/BE Signal/BEsignal.htm).....	7
Figura 4.2 – Tabela 1.1 Fonte: Webster J. G., Editor: Medical Instrumentation, Application and Designer. Wiley, 1998.	8
Figura 4.3 – Diagrama em blocos para aquisição de sinais bioelétricos	8
Figura 4.4 – Potencial de Ação.(Fonte: PARREIRA, 2006)	10
Figura 4.5– Localização dos ritmos cerebrais (Extraído de PARREIRA, 2006).....	11
Figura 4.6 – Diferentes tipos de ondas eletroencefalográficas normais. Fonte: http://www.infoescola.com/neurologia/ritmo-cerebral/	12
Figura 4.7 – Forma de Onda Alfa. (Fonte: http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/08/a-onda-alfa)	12
Figura 4.8 – Forma de Onda Beta. (Fonte: http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/09/a-onda-beta)	13
Figura 4.9 – Forma de Onda Teta. (Fonte: http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/10/a-onda-teta)	13
Figura 4.10 – Forma de Ondas Delta. (Fonte: http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/10/a-onda-delta)	13
Figura 4.11 – Simbolização de série temporal.	15
Figura 4.12 – Exemplo aleatório de distribuição de frequências absolutas da série temporal	16
Figura 4.13 – Exemplo da transformação dos símbolos com $n=5$	17
Figura 4.14 – Estimativa de atraso entre sinais de eletromiograma e estabilograma com a técnica tradicional de correlação cruzada, Atraso médio estimado de 1,11 s. A linha contínua indica o limiar de significância estatística de $p < 0,05$ (Fonte: MELLO et al., 2006).....	18

Figura 4.15 – Estimativa de atraso entre sinais de eletromiograma e estabilograma com a técnica de dinâmica simbólica. Atraso médio estimado de 1,16 s. A linha contínua indica o limiar de significância estatística de $p < 0,05$ (Fonte: MELLO et al., 2006).....	19
Figura 4.16 - Modelo rede neural e seus elementos (PORTILHO, et al. 2006).....	25
Figura 4.17 – Esquema de controle de padrão industrial (PORTILHO, et al. 2006).....	26
Figura 4.18 – Indivíduos de uma população e a sua correspondente roleta de seleção.	35
Figura 4.19 – Exemplo de mutação.....	36
Figura 4.20 – Um exemplo de crossover de um ponto. (a) dois indivíduos são escolhidos. (b) um ponto de crossover é escolhido. (c) são recombinadas as características, gerando dois novos indivíduos.....	37
Figura 5.1 – Simbolização de série temporal.	41
Figura 5.2 – Distribuição de frequência absoluta para uma série temporal periódica.	42
Figura 5.3 – Rede Neural Perceptron Multicamadas.....	45
Figura 5.4 – Gráfico de uma possível superfície de erro mostrando os mínimos locais e o mínimo global.....	47
Figura 6.1 – Diagrama de blocos do sistema, mostrando suas possibilidades de navegação.....	51
Figura 6.2 – Tela principal do Sistema.....	52
Figura 6.3 – Tela de Simbolização dos Sinais.....	52
Figura 6.4 – Tela do Simbolizador Automático.	53
Figura 6.5 – Tela do Otimizador Algoritmo Genético.	53
Figura 6.6 – Tela para separar os eventos e salvá-los.	54
Figura 6.7 – Tela principal do Classificar RNA.....	54
Figura 6.8 – Tela de ajuste dos parâmetros da Rede neural Artificial.....	55
Figura 6.9 – Tela Conjunto de Treinamento.....	55
Figura 6.10 – Tela de Construção e Treinamento da Rede Neural Artificial.....	56
Figura 6.11 – Tela do Classificador.....	56
Figura 6.12 – Figura de um sinal caótico.	57
Figura 6.13 – Figura de um sinal periódico.....	57

Figura 6.14 – Frequência relativa de sinal caótico utilizando parâmetros de simbolização aleatórios.....	59
Figura 6.15 – Frequência relativa de dois sinais periódicos utilizando parâmetros de simbolização aleatórios.	60
Figura 6.16 – Frequência relativa do sinal Caótico após a simbolização com os valores dos parâmetros otimizados pelo Algoritmo Genético.	61
Figura 6.17 – Frequência relativa dos Sinais Periódicos após a simbolização com os valores dos parâmetros otimizados pelo Algoritmo Genético.	61
Figura 6.18 – Tela do programa demonstrando o sinal Periódico normal e abaixo o sinal Periódico adicionado de ruído aleatório.....	62
Figura 6.19 – Frequência relativa de um sinal Periódico com ruído após a simbolização.	62
Figura 6.20 – Frequência relativa de um sinal Caótico com ruído após a simbolização.	63
Figura 6.21 – Separação dos eventos para formação da Base de Dados.	64
Figura 6.22 – Frequência Relativa dos sinais sem a presença de eventos epiléticos utilizando os valores da Tabela 6.3.	64
Figura 6.23 – Frequência Relativa dos sinais com a presença de eventos epiléticos utilizando os valores da Tabela 6.3.	65
Figura 6.24 – Frequência relativa dos Sinais Normais após a simbolização utilizando os parâmetros da Tabela 6.5.	67
Figura 6.25 – Frequência relativa dos Sinais Epiléticos após a simbolização utilizando os parâmetros da Tabela 6.5.	67

LISTA DE SIGLAS

ω_{ij} - Peso da rede neural

d - Saída desejada apresentado durante o treinamento da rede neural

i_{ep} - Soma do valor da entropia dos sinais epiléticos

i_n - Soma do valor da entropia dos sinais normais

η - Taxa de aprendizagem da rede neural

$\mathcal{E}_{(\eta)}$ - Valor da função de custo

$\Delta\omega_{(\eta)}$ - Valor de ajuste da rede neural

$e_{(\eta)}$ - Valor do erro da rede neural

$e_{(n)}$ - Valor do erro da rede neural no treinamento

$e_{i(\eta)}$ - Valor do estimo da rede neural

y - Saída real da rede neural

AGs – Algoritmos Genéticos

AR – Auto-Regressivo

C – Central

ECG - Eletrocardiograma

EEG – Eletroencefalograma

EMG – Eletromiograma

ENG - Eletroneurograma

EOG – Electrooculograma

ERG – Eletroretinograma

F – Frontal

FN – Falso negativo

FP – Falso positivo

Fp – Pólo frontal

Hz – Hertz

K+ – CDs Potássio

MLP – Multi-Layer Perceptron

Na⁺ – Sódio

N_{obs} – Número de colunas do histograma diferente de zero.

O – Occipital

OMS – Organização Mundial de Saúde

P – Parental

$P_{i,L}$ – Frequência relativa dos valores de cada coluna do histograma

PJ – Passo da Janela

PPS – Potencial pós-sináptico

RNA – Rede Neural Artificial

STFT – Transformada breve Fourier

TA – Tamanho do Alfabeto

TEO – Teager Energy Operation

TEP – Tomografia por emissão de pósitrons

TJ – Tamanho da Janela

TP – Tamanho da Palavra

TRME – Tomografia por ressonância magnética estrutural

VN – Verdadeiro negativo

VP – Verdadeiro positivo

VPN – Valor preditivo negativo

VPP – Valor preditivo positivo

WT – Wavelets

μ V – Microvolts

1. INTRODUÇÃO

O grande desafio do homem nos últimos tempos vem sendo as descobertas da cura de patologias que levam à morte ou comprometem a qualidade de vida. Nesse sentido, juntamente com os grandes males, surgem as necessidades de inovações para detecção rápida e eficiente destes males, pois com agilidade e precocidade na definição do diagnóstico, certamente o tratamento se mostra mais eficaz levando à obtenção de resultados mais satisfatórios.

O avanço da medicina ocorre paralelo ao da tecnologia, que vem se desenvolvendo rapidamente, e suas inovações vem trazendo resultados surpreendentes, e ao mesmo tempo, modificando e ampliando as formas de prevenção, tratamento e cura das mais variadas doenças. Portanto, esta pesquisa propõe um novo método para auxiliar na identificação e tratamento de um mal que afeta a vida de milhares de pessoas em todo o mundo, denominado de epilepsia.

Nosso cérebro é constituído por neurônios, que estão interligados entre si trocando informações através de impulsos elétricos, eventualmente há uma atividade elétrica anormal entre esses neurônios, onde o resultado visível desta anormalidade são as crises intermitentes. Quando essas crises acontecem freqüentemente em uma pessoa, diz-se que ela possui epilepsia. Uma crise é uma situação temporária, onde no caso da crise epiléptica o cérebro é tomado por descargas elétricas rápidas e fortes. Estas crises são caracterizadas principalmente por ataques onde a pessoa fica inconsciente, não se lembrando do acontecimento e também por dormência temporária nos membros.

Atualmente, sabe-se que a epilepsia é causada por células cerebrais que descarregam energia elétrica quando não deveriam, mas não se sabe exatamente a causa desse fenômeno. Em pesquisas diagnosticou-se que pessoas passaram a adquirir a doença após um: trauma craniano, problemas de saúde durante a gravidez, efeito tardio de doenças na infância como o sarampo, deficiências nutricionais e tumores cerebrais. (PSIQUIATRIA GERAL, 2010)

De acordo com a Organização Mundial de Saúde (OMS), países em desenvolvimento possuem maior incidência de indivíduos com epilepsia, e isso se deve ao fato da precariedade da saúde pública e à falta de diagnóstico e tratamento precoce. Nos países em desenvolvimento segundo a OMS os números chegam a 120-199 a cada 100 mil habitantes, enquanto em países desenvolvidos esse número está em cerca de 40-50 a cada 100 mil habitantes. Segundo a Assistência à Saúde de Pacientes com Epilepsia (ASPE), estima-se que existam no Brasil cerca de 3 milhões de pessoas com epilepsia e a cada hora são diagnosticadas 11 pessoas com a doença.

A epilepsia é a desordem neurológica mais comum na população mundial, podendo ser detectada e avaliada através do exame de eletroencefalografia (EEG), onde as características principais para existência de epilepsia são a presença de espículas agudas no registro temporal dos sinais do exame.

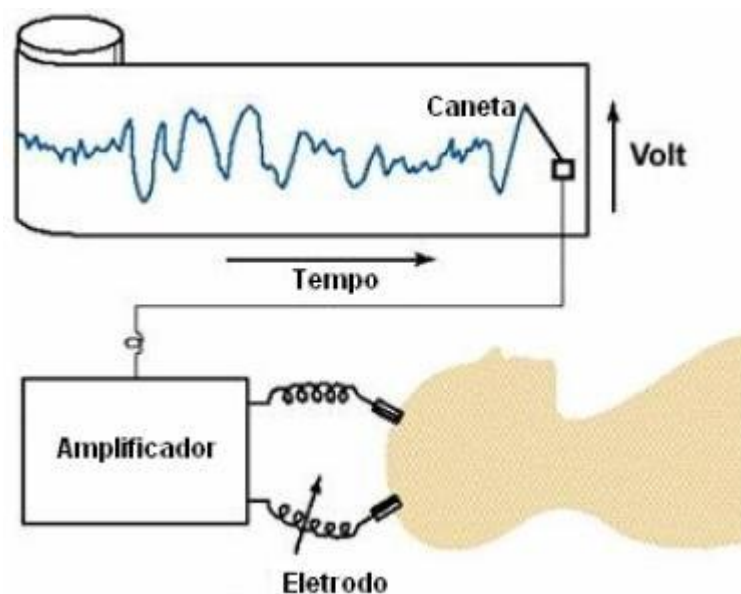


Figura 1-1 – Registro de atividade de eletroencefalográfica. Fonte: http://www.cerebromente.org.br/n03/tecnologia/eeg_p.htm

O sinal de EEG retrata as atividades em potencial elétrico captado na região do encéfalo (Figura 1.1), gerando assim um diagrama onde se pode avaliar a extensão, a dinâmica, localização e distribuição da atividade cerebral.

Os resultados provenientes do exame de EEG para detecção de presença ou não de sinais epiléticos são feitos visualmente pelo profissional, tomando muito tempo para a avaliação dos resultados. Na maioria dos exames, que têm duração média de trinta

minutos, podem conter até quarenta páginas de registros gráficos, não sendo às vezes suficientes para um diagnóstico conclusivo, podendo haver necessidade de se alongar a duração do exame em duas a três horas. Em casos mais complexos, pode-se haver necessidade de uma observação por meio de filmagem de dois a três dias para um laudo conclusivo.

Haja em vista a dificuldade de se obter um laudo conclusivo com agilidade devido à extensão do exame visual dos registros, esse trabalho propõe um *software* capaz de resolver esta questão de forma automática dos sinais de EEG. Utilizando-se o método de simbolização otimizada por algoritmos genéticos de séries temporais e o treinamento de uma rede neural artificial onde serão classificados sinais epiléticos e não epiléticos, podendo vislumbrar toda dinâmica das atividades cerebrais de pessoas com epilepsia ou não. Ressalta-se que o *software* proposto como resultado do estudo realizado conta com uma interface gráfica amigável para ser utilizada pelo usuário final. Pretende-se, dessa forma, contribuir para o diagnóstico mais ágil e preciso para o paciente, bem como, para a redução do seu custo.

2. OBJETIVO E METAS

2.1. OBJETIVO

O objetivo principal deste trabalho foi desenvolver um método capaz de avaliar os sinais de EEG, utilizando dinâmica simbólica para a descrição das séries temporais, classificando esses resultados em uma rede neural artificial e finalmente implementar a metodologia em um *software* capaz de gerar um resultado automático, que possa ser facilmente avaliado pelo usuário final, de modo a ser testado e validado experimentalmente por profissionais da área de eletroencefalografia.

2.2. METAS

As metas alcançadas durante a realização deste trabalho foram:

Otimização por Algoritmos Genéticos, a descrição por meio de dinâmica simbólica de sinais de EEG de pacientes comprovadamente epiléticos e não epiléticos;

Desenvolvimento de um sistema para identificar automaticamente os sinais oriundos de pacientes epiléticos utilizando redes neurais artificiais;

Verificação da eficácia do protótipo, por meio da utilização de registros de EEG obtidos junto a profissionais da área de eletroencefalografia e neurologistas;

Desenvolvimento de uma interface gráfica que possa ser bem visualizada e entendida pelo usuário final, de modo a tornar os laudos do exame de EEG mais rápidos e eficientes, a ponto de oferecer resultados concisos e confiáveis.

3. ORGANIZAÇÃO

A estrutura deste trabalho está organizada em 8 capítulos, conforme descrição:

O Capítulo 1 descreve a introdução deste trabalho.

Os Capítulos 2 e 3 apresentam os objetivos, as metas e a organização geral.

O Capítulo 4 descreve a revisão bibliográfica da literatura científica na área em questão. Destacando-se a Dinâmica e Análise de Sinais Bioelétricos, Dinâmica Simbólica, Análise de sinais caóticos, Redes neurais artificiais e os Algoritmos genéticos.

Destaca-se no Capítulo 5 a metodologia empregada neste trabalho, com os passos empregados para chegar aos objetivos, descrevendo as formas para obtenção dos dados de EEG, a Simbolização das séries, a classificação dos sinais, os índices estatísticos de desempenho e a modelagem do *software*.

A Apresentação dos resultados obtidos com os métodos descritos comparando com outros resultados da literatura são apresentados no Capítulo 6.

O Capítulo 7 relata a conclusão e discussão sobre os resultados obtidos, discutindo as formas utilizadas e comparando com trabalhos que embasaram essa dissertação.

Para finalizar, o Capítulo 8 apresenta os trabalhos futuros os quais poderão aprimorar e utilizar as técnicas descritas neste trabalho.

4. REVISÃO BIBLIOGRAFICA

4.1. DINÂMICA E ANÁLISE DE SINAIS BIOELÉTRICOS

Os sinais bioelétricos são resultados de diversas movimentações celulares, podendo ser de características químicas, mecânicas e elétricas. Todas estas movimentações são geradas através das células nervosas e musculares, gerando enfim os sinais bioelétricos, que não se constituem de forma individual, mas por um grupo de células.

Sinais Bioelétricos estão associados às diversas variáveis fisiológicas (função, disfunção ou funcionamento de órgãos e tecidos do corpo humano), refletindo as diversas formas como são conduzidas as informações dentro do organismo. As ações geradas pelo acúmulo de efeitos produzidos pelas células propagam-se pelas células e tecidos vizinhos, criando-se um campo elétrico que é propagado pelos vários tecidos do corpo. Portanto podemos medir uma série de atividades elétricas em nosso corpo utilizando como, por exemplo, eletrodos posicionados na pele no caso de atividades musculares e neuronais (GUYTON, 1997).

As ondas elétricas que se propagam através das membranas celulares são chamadas de Potencial de Ação, conhecidas também como sinais bioelétricos, muito importantes para os seres vivos, pois transportam informações entre os tecidos do corpo. Diversos tipos de células podem-se utilizar deste potencial de ação para se comunicarem. Estes sinais são medidos com a ajuda de sistemas computadorizados onde passam por uma análise utilizando processamento digital (CRUZ, 2007).

Eletroencefalograma – EEG (ondas elétricas do cérebro);

Eletroneurograma – ENG (velocidade de condução dos nervos);

Eletrooculograma – EOG (movimentos dos olhos);

Eletroretinograma – ERG (transdutor da luz sinais elétricos ao cérebro);

Eletromiograma – EMG (tensão muscular);

Eletrocardiograma – ECG (ondas elétricas do coração);

Pneumograma – (movimentos respiratórios e a oxigenação do sangue).

Devido à variedade de tipos de sinais bioelétricos que se podem captar, vê-se que suas características são bem diferentes dependendo dos locais do corpo em que são registrados, na Figura 4.1 temos o exemplo de registro como o ECG; EEG; EMG e o EOG:

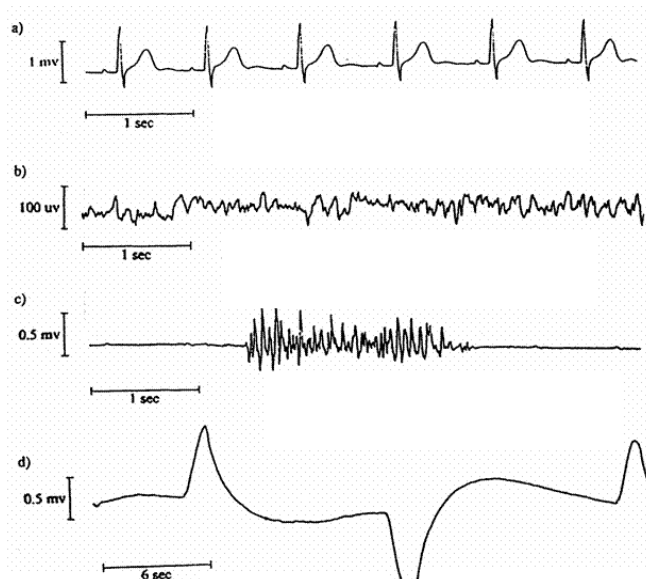


Figura 4-1 – Sinais bioelétricos retirados de diferentes partes do corpo: a) ECG; b) EEG; c) EMG e d) EOG. (Fonte: IEM0700 Biomedical Electronics, <http://www.elin.ttu.ee/mesel/Study/Courses/Biomedel/Content/BioElect/BESignal/BESignal.htm>)

Na Tabela 4.1 pode-se observar a variação de medida entre os sinais e diferenças entre os biosinais, observa-se que a amplitude dos sinais bioelétricos são da ordem microvolts ($1 \mu\text{V} = 10^{-6} \text{V}$) (CRUZ, 2007).

Parâmetros	Faixa de variação	Faixa de frequência, Hz	Método de transdução
Eletrocardiografia (ECG)	0,5 – 4 mV	0,01 – 250	Eletrodos na pele
Eletroencefalografia (EEG)	5 – 300 μ V	dc-150	Eletrodos no couro cabeludo
Eletromiografia (EMG)	0,1 – 5 μ V	dc-10 ⁴	Eletrodos tipo agulha
Eletrocorticografia (ECOG)	10 – 5000 μ V	dc-150	Eletrodos de superfície
Fonocardiografia (PCG)	Abrang. Dinâmica 8 dB, limiar cerca de 100 μ Pa	8 – 2000	Microfone

Resposta galvânica da pele (GSR)	1 – 500 K Ω	0,01 – 1	Eletrodos de pele
Potenciais neurais	0,01 – 3 mV	dc-10 ⁴	Eletrodos tipo agulha ou de superfície

Tabela 4.1 - Fonte: Webster J. G., Editor: Medical Instrumentation, Application and Designer. Wiley, 1998.

4.2. AQUISIÇÃO DOS SINAIS BIOELÉTRICOS

A seqüência de uma aquisição de sinais bioelétricos pode ser representada pelo diagrama em blocos mostrado na Figura 4.2.

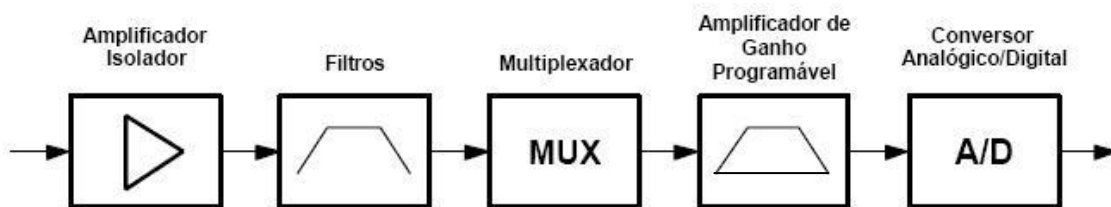


Figura 4-2 – Diagrama em blocos para aquisição de sinais bioelétricos

Para a captação dos sinais bioelétricos são utilizados eletrodos, que são pequenas peças de metal fixados em locais específicos do corpo do paciente, conectados a fios que transmitem os sinais captados para um amplificador .

Devido a pequena amplitude dos sinais captados pelos eletrodos, o amplificador tem a função de amplificar estes sinais para que possam ser observados e registrados. Em seguida, passam por um processo de filtragem onde sinais espúrios indesejáveis inerentes ao processo de medição, podendo interferir no sinal de interesse são atenuados (CRUZ, 2007). O amplificador é construído para funcionar como um isolador elétrico, evitando que o paciente não corra riscos de sofrer descargas elétricas (BORN, 2000).

Os sinais bioelétricos possuem uma grande diversidade de amplitudes, com isso os amplificadores dão ganho a estes sinais, tendo assim um nível de ganho de sinal controlado, sendo essa função dos amplificadores programáveis (BORN, 2000).

Finalmente, os conversores transformam os sinais analógicos em sinais digitais através de amostragens, os principais parâmetros para conversão de sinais analógicos para sinais digitais são os seguintes (TAVARES, 1997):

- A taxa ou velocidade de conversão, que define a quantidade de amostras tomadas do sinal analógico por unidade de tempo, e geralmente é expressa em “amostras/s” ou em “Hz”.
- A taxa de conversão escolhida para uma determinada aplicação deve levar em conta o critério de Nyquist, que determina que taxa de conversão deva ser no mínimo duas vezes maior que a frequência máxima do sinal desejado, a fim de evitar perda de informação (*aliasing*).
- Resolução, que define qual é o mínimo salto de tensão detectável no sinal de entrada, e é expressa em número de bits. Por exemplo, se um determinado conversor Analógico/Digital apresenta uma resolução de 8 bits, é possível discriminar 256 níveis diferentes de tensão.

4.3. INTRODUÇÃO AO SINAL DE EEG

As pesquisas indicam que os neurônios são células responsáveis pelo armazenamento e transmissão de informação dentro do sistema nervoso central, e toda transmissão ou comunicação neural envolve certo fluxo de íons. Como os íons possuem cargas elétricas, é possível registrar potenciais elétricos dentro e fora da parede celular (KUGLER, 2003).

Nessa linha de raciocínio, estes sinais são captados através de eletrodos colocados na superfície do couro cabeludo, posicionados nas regiões frontal, central, parietal e occipital. Os diferentes registros de potenciais elétricos que são gerados são ocasionados pelas diferentes movimentações celulares de forma simétrica e ordenada de certas classes de células do córtex cerebral, sendo que para se gerar estes potenciais é necessário que estes sejam disparados por um número significativo de neurônios (BUTTUN, 2000).

Conforme Button (2000), os potenciais elétricos ou biopotenciais encefálicos podem ser diferenciados em dois tipos:

Potencial de Ação: decorrente da transmissão dos impulsos elétricos pelos axônios. Não tem muita contribuição para o registro de EEG, uma vez que ocorrem de forma assíncrona e em direções diversas ortogonais à superfície;

Potencial Pós-Sináptico (PPS): conforme as entradas excitatórias (dendritos com polaridade negativa em relação ao corpo celular e o potencial negativo registrado na superfície cortical) e inibitórias (dendritos com polaridade positiva em relação ao corpo celular; potencial positivo registrado na superfície cortical) alcançam os dendritos de cada neurônio, “flui corrente” para os dendritos ou para os corpos celulares, respectivamente, definindo dipolos elétricos. Os dipolos mudam de intensidade e sentido, produzindo flutuações ondulares no potencial elétrico resultante no volume condutor. O potencial registrado na superfície cortical depende da polaridade, orientação e localização do PPS em relação ao eletrodo da medida.

No interior das superfícies do escalpo (EEG) ou na superfície do córtex (ECoG) podemos registrar diferenças de potencial que ocorrem no interior do tecido neural, onde os fluxos de correntes extracelulares são produzidos pelos potenciais pós-sinápticos e pelos potenciais de ação dos axônios.

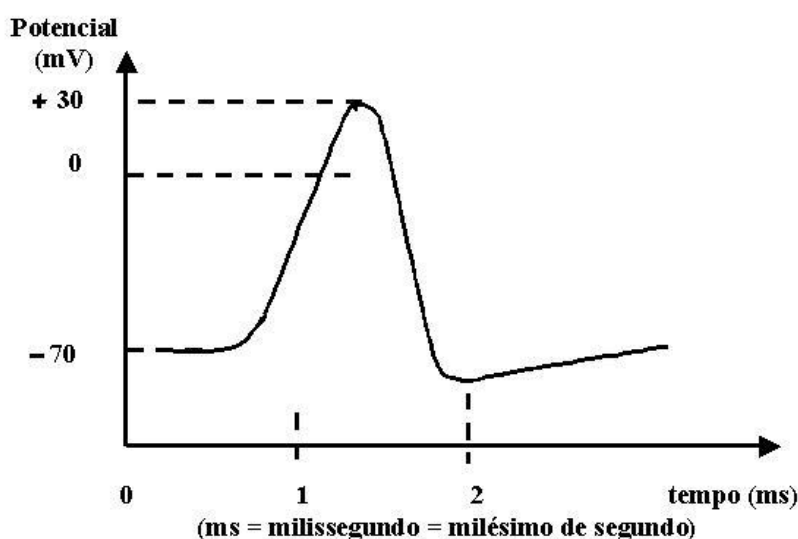


Figura 4-3 – Potencial de Ação. (Fonte: PARREIRA, 2006)

Segundo Guyton (1997), as intensidades das ondas cerebrais registradas do escalpo variam entre $0 \mu\text{V}$ e $200 \mu\text{V}$ e sua faixa de frequência vão de 0,3 Hz a 100 Hz. A intensidade destes sinais juntamente com seus padrões pode significar alguma atividade encefálica como o sono, psicoses e doenças nervosas como a epilepsia. Por sua vez, os registros ondulares das atividades elétricas são chamados de ondas cerebrais e a totalidade do registro é o eletroencefalograma (EEG). Logo, os sinais de EEG são um reflexo de atividades bioelétricas dos neurônios no cérebro (KUGLER, 2003 *apud* GUYTON, 1992).

Portanto, as ondas registradas através do sinal de EEG, são bastante irregulares, porém algumas patologias possuem características de traçados bastante padronizados, como a epilepsia. Em indivíduos normais podem aparecer como ondas classificadas como Alfa, Beta, Teta e Delta (KUGLER, 2003, *apud* GUYTON, 1992).

4.3.1. Morfologia do Sinal EEG

As formas como são caracterizados os sinais de EEG estão intrinsecamente relacionados com o posicionamento dos eletrodos no escalpo. Ou seja, dependendo do local onde está posicionado, irão registrar ondas típicas daquela região, conforme Figura 4.4.

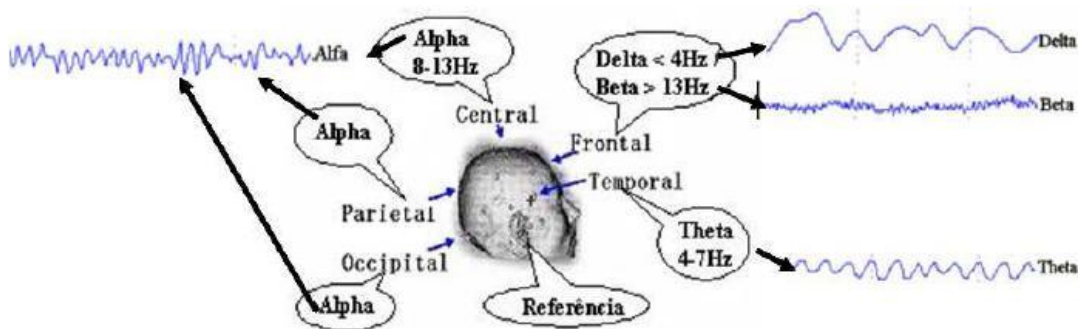


Figura 4-4– Localização dos ritmos cerebrais (Extraído de PARREIRA, 2006).

Os sinais possuem particularidades, não só pelas suas características visuais de registro, mas também por se referirem ao tipo de captação que podem ser classificadas como ondas Alfa, Beta, Teta e Delta, conforme mostradas na Figura 4.5.

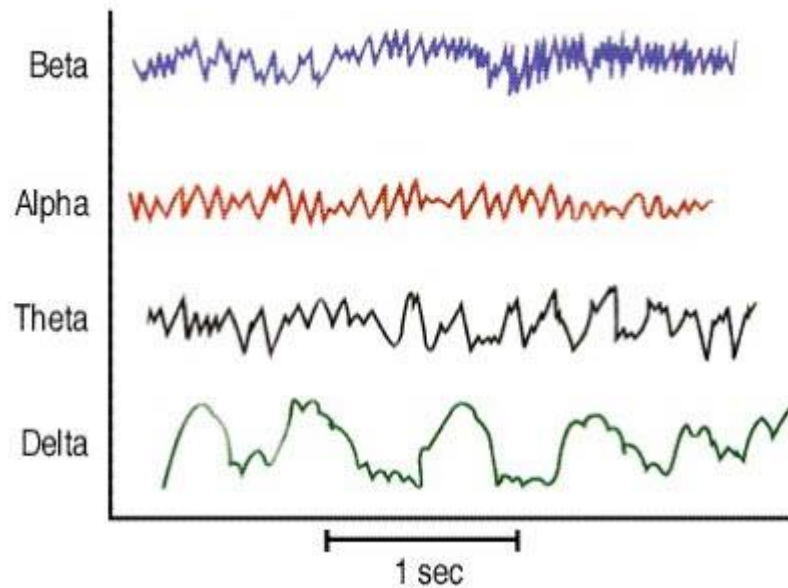


Figura 4-5 – Diferentes tipos de ondas eletroencefalográficas normais. Fonte:
<http://www.infoescola.com/neurologia/ritmo-cerebral/>

Segundo o pesquisador Parreira (2006), ondas Alfa são as ondas mais importantes nos estudos sobre o estado de vigília, embora haja variação de sua quantidade entre os indivíduos normais. Elas ocupam uma frequência entre 8 a 13 Hz, e geralmente são encontrados na maioria dos indivíduos que estão em estado relaxado, consciente e com os olhos fechados, diminuindo sua frequência quando o indivíduo abre os olhos, fica em estado de sonolência ou sono, além de serem mais comuns em adultos.

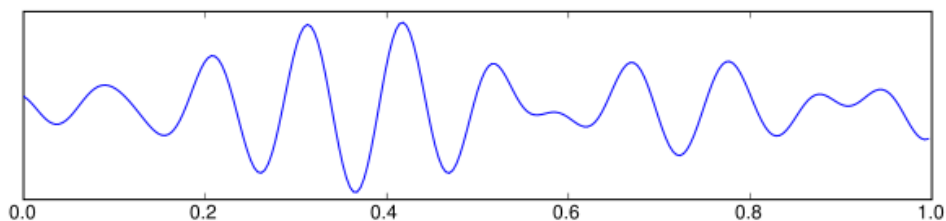


Figura 4-6 – Forma de Onda Alfa. (Fonte:
<http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/08/a-onda-alfa>)

As ondas Beta estão presentes na faixa de frequência acima de 13 Hz. Alguns tipos de drogas podem aumentar esse tipo de onda e a principal característica é sua presença em todas as idades. É um tipo de onda características de indivíduos em estado de ansiedade e tensão.

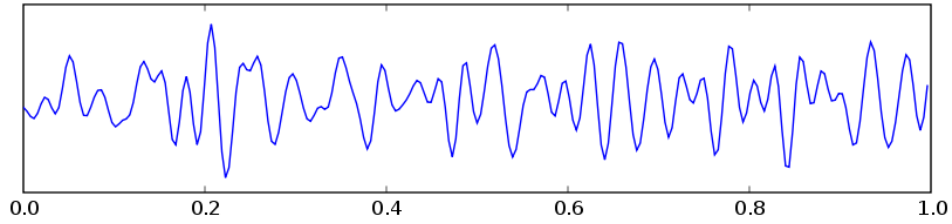


Figura 4-7 – Forma de Onda Beta. (Fonte: <http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/09/a-onda-beta>)

As ondas Teta possuem frequência entre 4 a 7 Hz, pesquisas indicam que é bastante presente em crianças e dificilmente se manifestam em pessoas com uma idade mais avançada, sendo considerado um sinal e estado característico de meditação e criatividade. (PARREIRA, 2006 *apud* DENIS, 1999)

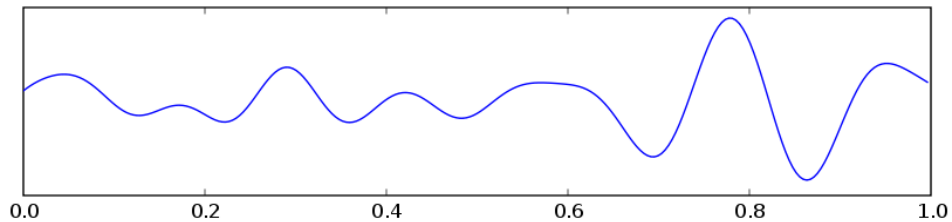


Figura 4-8 – Forma de Onda Teta. (Fonte: <http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/10/a-onda-teta>)

As ondas Delta possuem frequência entre 0,5 a 4 Hz no sinal EEG. É um tipo de onda que é encontrada em bebês, crianças, em adultos com sono profundo, sendo anormais em adultos acordados.

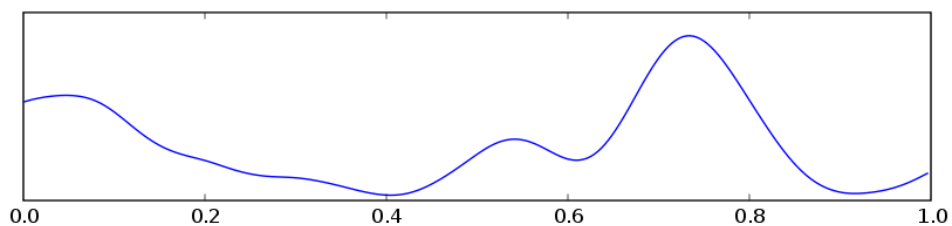


Figura 4-9 – Forma de Ondas Delta. (Fonte: <http://blogues.cienciahoje.pt/index.php/Impulsos/2011/02/10/a-onda-delta>)

As variações destas ondas fora do ritmo ou estado aos quais cada uma pertence geralmente caracterizam alguma manifestação de patologia cerebral. As diversas regiões e seus ritmos estão descritos na Tabela 4.2 abaixo.


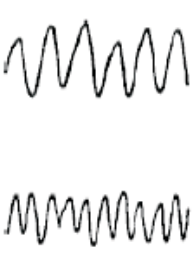
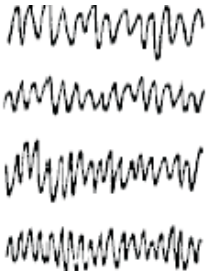
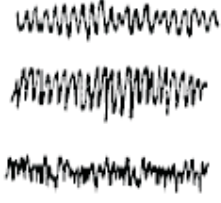
RITMO	DELTA	TETA	ALFA	BETA
Componente em frequência	< 4 Hz	4 a 7 Hz	8 a 13 Hz	> 13 Hz
Amplitude	100 μV	Criança: 20 μV Adulto: 10 μV	Bebê: 20 μV Criança: 75 μV Adulto: 50 μV	10 a 20 μV
Principal área do escalpo	Frontal	Temporal	Occipital Parietal	Frontal
Condição do indivíduo	Sono profundo (adulto)	Sonolência	Repouso olhos fechados	Repouso olhos abertos
Forma de onda				

Tabela 4.2– Ritmo das principais ondas cerebrais (Extraído de PARREIRA, 2006).

O sinal de EEG é bastante utilizado para diagnosticar distúrbios e danos no que se refere ao funcionamento do cérebro. Como o sinal é bastante intenso com uma gama enorme de informações torna-se indispensável o desenvolvimento de técnicas que auxiliem na identificação automática de patologias para somar e auxiliar os especialistas.

4.4. DINÂMICA SIMBÓLICA

4.4.1. Introdução

A Dinâmica Simbólica freqüentemente estudada no contexto da dinâmica de sistemas, baseia-se na discretização de sinais temporais em seqüências de símbolos que representam os estados do sistema.

Segundo Daw (2003), descreve-se a utilização das séries temporais experimentais, onde um sinal temporal é representado por uma série temporal discreta $T = \{T_1, T_2 \dots T_n\}$, cada medida equivale a um estado, cada estado pode ser representado por um símbolo n que pertence a um alfabeto finito. As séries temporais passam por um processo de conversão onde os processos de saída acabam assumindo valores discretos e finitos, além de receberem valores utilizando o método de transformação. Por exemplo, utilizando um alfabeto binário ($n=2$), os valores são distribuídos de acordo com um valor que define os símbolos “0” ou “1”, segundo a medida do dado limiar, assim representado pela Figura 4.10.



Figura 4-10 – Simbolização de série temporal.

Após simbolizar os dados do sinal, são definidas seqüências de símbolos de tamanho m , onde são realizadas as freqüências relativas e calculadas para seqüência completa obtida da discretização de todas as possíveis seqüências de símbolos possíveis de tamanho m (ver Tabela 4.3).

São representados o comprimento m por um equivalente decimal, por exemplo, uma sequência de 011_2 (base binária) passa a ser representada por uma base decimal 3_{10} . Assim é possível representar de forma estatística a distribuição das sequências simbólicas, podendo ser visto através de um histograma como na Figura 4.11.

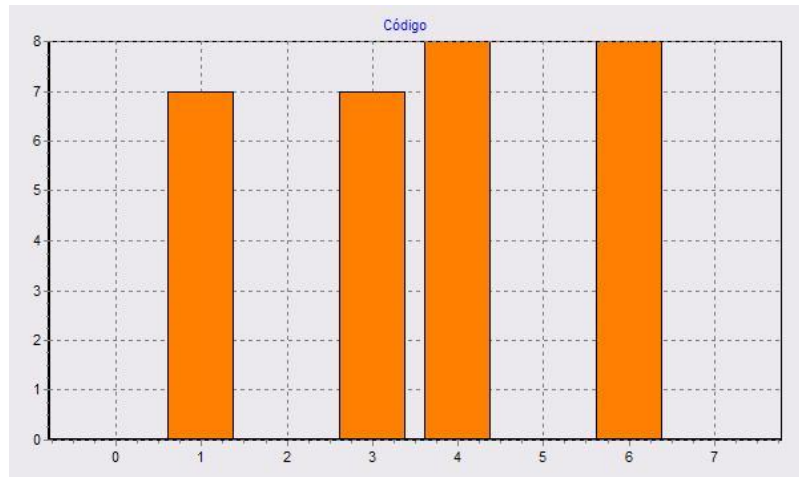


Figura 4-11 – Exemplo aleatório de distribuição de frequências absolutas da série temporal

Seqüência	Código	Frequência
000	0	0
001	1	7
010	2	0
011	3	7
100	4	8
101	5	0
110	6	8
111	7	0

Tabela 4.3 – Exemplo aleatório da frequência das sequências

Pode-se trabalhar também com diversos valores de n , neste caso a simbolização terá mais equivalentes limiares, passa a não ser representados apenas por “0” e “1”, mas o tamanho ganha diversos outros valores representativos como o exemplo de $n=5$ mostrado na Figura 4.12.

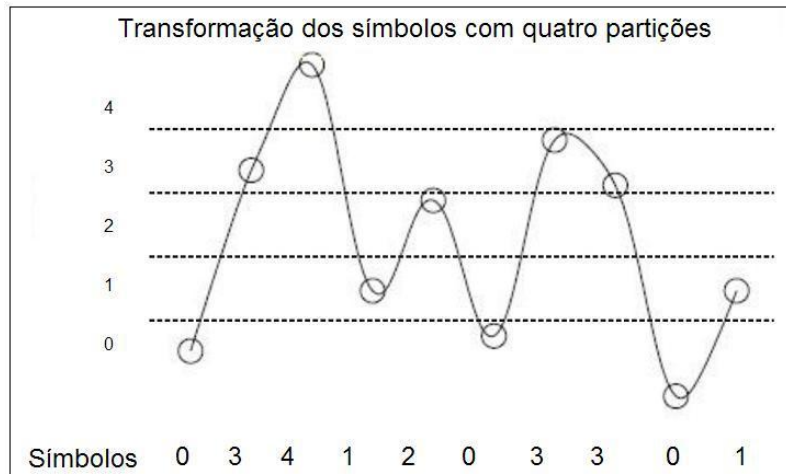


Figura 4-12 – Exemplo da transformação dos símbolos com $n=5$.

Voss *et al.*, (1996) pesquisaram a variabilidade da frequência cardíaca entre pessoas saudáveis e pessoas que sofreram infarto utilizando as técnicas tradicionais e a Dinâmica Simbólica. Esses sinais foram obtidos por meio de exame de eletrocardiograma, onde foram codificados e transformados em quatro símbolos. Os pesquisadores observaram que a Dinâmica Simbólica é mais eficiente para diagnosticar as pessoas com problemas cardíacos quando comparada a outras técnicas tradicionais.

A Dinâmica Simbólica foi utilizada em sinal de volume respiratório tendo como objetivo definir as variabilidades utilizando o mínimo de parâmetros possíveis (CAMINAL *et al.*, 2004). As amostras da série $RN = (\text{tempo inspiratório do ciclo } n / \text{período respiratório do ciclo } n)$ foram todos codificadas em três símbolos diferentes de acordo com sua amplitude:

$$S_n = \begin{cases} 0, & \text{se } R_n < L_1, \\ 1, & \text{se } L_1 < R_n < L_2 \\ 2, & \text{se } L_2 < R_n \end{cases}$$

Como funções da amplitude média do sinal de respiração e de um parâmetro adimensional α , foram escolhidos os limiares L_1 e L_2 . Eles analisaram palavras com três símbolos consecutivos, onde teriam uma combinação possível de 27 palavras diferentes. Durante o experimento utilizaram três parâmetros: o parâmetro α , o limiar de ocorrência das palavras e o grau de superposição em palavras consecutivas, onde são analisadas as

diferentes palavras que possuem símbolos iguais nas mesmas posições (ex.: 001 e 101 notam-se que são palavras diferentes, porém possuem dois símbolos nas mesmas posições iguais, diferentemente de 101 e 010 que não possuem nenhuma superposição). A Dinâmica Simbólica foi capaz de resolver esse problema classificando os sinais de respiração em dois grupos, baixa e alta variabilidade, com a mesma eficácia dos critérios clínicos, porém de uma forma automática e concisa (CAMINAL *et al.*, 2004).

Mello *et al* (2006), desenvolveram uma aplicação utilizando Dinâmica Simbólica, onde o objetivo principal era quantificar a dinâmica no equilíbrio postural ortostático. Para tanto, foram coletados para esse experimento os sinais de eletromiograma (EMG) e estabilograma onde foram codificadas as séries da Dinâmica Simbólica de acordo com as derivadas discretas de cada momento no espaço dos sinais, sendo os valores das derivadas dos sinais positivos recebendo valor 1 e para as derivadas dos sinais negativos 0. Para calcular a correlação cruzada entre os dois sinais, foram calculadas e comparadas as coincidências entre as séries da Dinâmica Simbólica, onde é denominada “Dinâmica Simbólica Conjunta” e que obteve resultados semelhantes aos da correlação cruzada. Portanto, a Dinâmica Simbólica obteve os mesmos resultados e desempenho das técnicas tradicionais de correlação cruzada, com uma vantagem da Dinâmica Simbólica resolver os problemas para o uso em sinais não linearmente correlacionados.

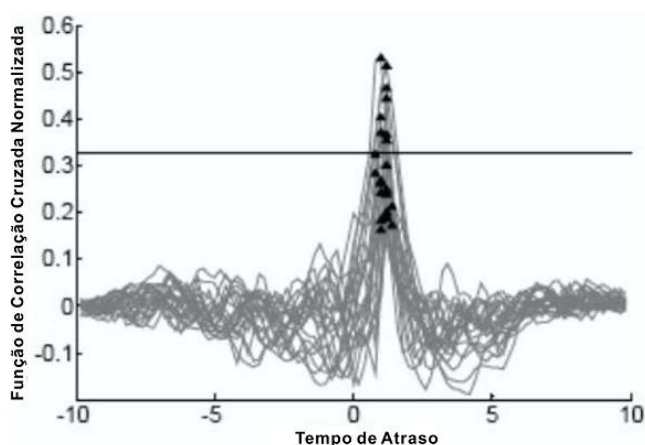


Figura 4-13 – Estimativa de atraso entre sinais de eletromiograma e estabilograma com a técnica tradicional de correlação cruzada, Atraso médio estimado de 1,11 s. A linha contínua indica o limiar de significância estatística de $p < 0,05$ (Fonte: MELLO *et al.*, 2006).

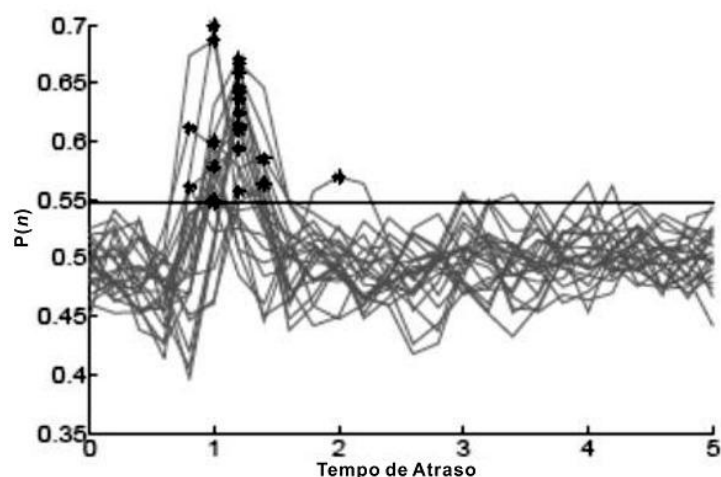


Figura 4-14 – Estimativa de atraso entre sinais de eletromiograma e estabilograma com a técnica de dinâmica simbólica. Atraso médio estimado de 1,16 s. A linha contínua indica o limiar de significância estatística de $p < 0,05$ (Fonte: MELLO *et al.*, 2006).

Outros pesquisadores fizeram comparações entre os principais parâmetros de pessoas com problemas crônicos do coração para avaliar seus riscos cardíacos. Com maior eficiência a Dinâmica Simbólica consegue resolver problemas maiores na detecção de anomalias na regulação autônoma cardiovascular, principalmente quando se comparam a outros parâmetros. Os pesquisadores fizeram uma relação da Dinâmica Simbólica que também foi aplicada em dados de sujeitos normais, com mudanças seletivas na regulação autonômicas cardiovasculares induzidas de forma experimental. Contudo, chegou-se a conclusão que a Dinâmica Simbólica pode auxiliar muito os profissionais da medicina, pois foi a Dinâmica Simbólica conseguiu detectar o aumento de riscos de eventos cardíacos, tornando assim uma ferramenta de auxílio clínico muito poderosa na prevenção. (MAESTRI *et al.*, 2007)

Por exemplo, a patente americana de número 6731972 descreve o uso de Dinâmica Simbólica para análise temporal de sinais cardiológicos, e sugere sua implementação em desfibriladores, marca-passos ou equipamentos afins.

4.4.2. Dinâmica de Sistemas

Os fatores não-lineares estão presentes em vários fenômenos naturais. A natureza tem como base sua própria vocação para os sistemas biológicos, que são sem dúvida a característica principal da vida. Desta forma surge a motivação para pesquisas relacionadas às características dinâmicas existentes nos sistemas naturais.

Os fenômenos naturais podem ser classificados em regulares onde são ditos periódicos ou irregulares. Um exemplo simples de fenômenos regulares são os sinais epileptiformes presentes no EEG durante eventos epiléticos, pode-se também observar comportamentos regulares normais em ciclos menstruais ou ritmos cardíacos.

O termo Doenças Dinâmicas tem relação com alguns tipo de ritmos anormais, que por sua vez, está inteiramente associado com as alterações rítmicas dos sistemas naturais, dando a entender que alguma anormalidade conduziu a um fator instável dentro do sistema, levando a uma profunda avaliação de como surgiu e como resolver essa anormalidade. A análise da dinâmica destes sistemas pode ser realizada tanto com a utilização de modelos matemáticos quanto também através da análise de sinais, onde podemos através destes aproximar de um resultado bem próximo dos fatores que levam a anormalidade desses sistemas.

Dentro deste contexto dos sistemas não-lineares existe uma grande possibilidade da existência do caos sobre esses sistemas. Seu comportamento é muito rico, contando com uma imensa variedade de órbitas periódicas instáveis que estão basicamente relacionadas a esses fatores. Uma característica muito importante que podemos observar no caos é sua sensibilidade às condições iniciais que envolvem diversos sistemas não-lineares, esse fator é denominado efeito borboleta. Tal efeito fez com que os estudos sobre o caos evoluíssem de forma significativa, onde pode ser explorado com bastante ênfase dentro dos sistemas caóticos, controlando-os com bastante eficácia observando as pequenas perturbações, o que proporciona uma grande flexibilidade ao sistema.

4.5. ANÁLISE DE SINAIS CAÓTICOS

4.5.1. Introdução

A informação está basicamente relacionada ao estudo dos sinais e sistemas, principalmente quando as informações se encontram na forma de dados digitais. Por exemplo, sinais como áudio, imagem e vídeo, além de outros como registros biométricos (eletrocardiogramas, eletroencefalogramas, bases genéticas etc.), geológicos (sísmicos), meteorológicos (eletromagnéticos, radares, sonares etc.) e financeiros (câmbio, mercado de valores etc.) são utilizados para reconhecimento de padrões, filtragem de bandas específicas e análise comportamental de previsões (CÂNDIDO, 2005).

Freqüentemente a análise de sinais e sistemas é feita com base na transformação da representação do sinal no domínio do tempo para o domínio da freqüência. Para tanto, o sinal deve satisfazer duas condições: linearidade e invariância no tempo. Desta forma, os sinais podem ser tratados matematicamente (COOLEY 1965, WINOGRAD 1978).

Alguns sistemas podem apresentar características de comportamentos não-lineares e variáveis, ou seja, podem apresentar características não-estacionárias, aleatórias ou de natureza caótica. Exemplos destes sinais são: Sinais biométricos, sinais de timbre, sinais de fala, turbulência, ruídos térmicos e quantização, sistemas de Lorenz, etc.

Nessa linha de raciocínio, outros sistemas são dificilmente tratáveis utilizando as técnicas tradicionais, tais como: transformadas de Fourier, transformadas Z e outras freqüentemente utilizadas para análise de sinais lineares.

No estudo de sistemas não-estacionários é necessário utilizar o “Janelamento”, que é basicamente observar o segmento de sinal presente dentro de um intervalo de tempo, onde se observa e localiza as regiões de interesse do sinal.

Os sistemas aleatórios ou estocásticos são imprevisíveis, mas com variações que podem possuir invariância estatística. A análise destes sistemas depende de uma descrição em termos de modelos estatísticos (CÂNDIDO, 2005).

Os sistemas caóticos podem ser modelados por equações diferenciais não-lineares. Os sistemas caóticos são facilmente confundidos com os sistemas aleatórios, entretanto possuem comportamento determinístico e alta sensibilidade a pequenas variações nas condições iniciais, onde qualquer fator de incerteza pode produzir grandes diferenças de resultados no decorrer de uma evolução temporal.

Existem várias técnicas que analisam os problemas relacionados com os sistemas não-lineares, tais como: algoritmos para análise de espaço de fase, análise por funções descritivas, análise com base na teoria de Lyapunov e filtros adaptativos. Além de contar com ferramentas matemáticas como: transformada breve de Fourier (STFT) e transformada Wavelets (WT). Estas ferramentas matemáticas trabalham como extratores de características e podem ser trabalhadas juntamente com análise polinomial, redes neurais, lógica Fuzzy e técnicas de correlação (espectro de potência), que podem atuar como classificadores. Existem vários métodos e técnicas para o tratamento de sinais, e várias pesquisas continuam sendo feitas para encontrar meios mais eficientes e específicos para aplicações nesta área (CÂNDIDO 2005).

4.6. REDES NEURAIAS ARTIFICIAIS PARA CLASSIFICAÇÃO

Com o avanço dos computadores, o homem vem sempre buscando desenvolver tecnologias capazes de serem independentes dos controles humanos. É através desta forma de pensar que surgiu a necessidade de se imitar o cérebro humano, cujo comportamento seja desenvolvido de acordo com seu próprio aprendizado, e que possa ser chamada de autônoma, cognitiva ou inteligente.

Os computadores trabalham de forma seqüencial, desenvolvendo tarefas que são programadas, já o cérebro humano trabalha de maneira paralela com imensas interconexões que facilitam e aumentam a capacidade de eficiência em tarefas e aprendizagens.

As redes neurais têm sido utilizadas para solucionar diversos problemas de engenharia como, por exemplo, em processamento digital de imagens, robótica e processamento digital de sinais, e também tem sido muito úteis para outras áreas como

geologia, física, biomédica e finanças. São técnicas computacionais desenvolvidas e inspiradas no funcionamento do cérebro humano, onde os neurônios são treinados para responder aos padrões de entrada, gerando uma saída, sendo utilizada para reconhecimento de padrões onde não se julga a necessidade de treinamento prévio.

Portanto, onde quer que existam problemas de previsão, classificação, análise de séries temporais e até mesmo de controle, deve ser avaliada a aplicação das Redes Neurais Artificiais. Nesse sentido, os objetivos das rede neural artificial são aprender com as experiências, e poder se adaptar em diversas situações. Estas características tornam-se grandes objetivos em suas pesquisas e implementações, sendo capazes cada vez mais a estarem sujeitas a diversos tipos de tolerância, como por exemplo, ruídos (HAYKIN, 1999).

O cérebro humano possui uma imensa capacidade de aprendizado se tornando um magnífico processador composto por neurônios onde todas as funções e movimentos estão relacionados aos neurônios. Estima-se que nosso cérebro possua mais de 100 bilhões de neurônios interligados através de uma sinapse, ela é capaz de transmitir estímulos através de diferentes concentrações de Na⁺ (sódio) e K⁺ (potássio). O resultado desta sinapse pode ser estendido a todo corpo como resultado desta interligação entre os neurônios. Esta grande concentração de neurônios interligados entre si é chamada de a famosa Rede Neural, que é capaz de processar e armazenar informações, essa complexa conectividade é responsável pelas características atribuídas à inteligência.

Pode-se dizer que as Redes Neurais Artificiais são sistemas simplificados de baixo nível das Redes Neurais Biológicas, pois se tenta imitar o funcionamento utilizando-se de técnicas matemáticas para seu funcionamento. Porém, enquanto uma Rede Neural Artificial em algumas aplicações podem utilizar de centenas de neurônios denominados de “nós”, as Redes Neurais Biológicas são compostas de bilhões de neurônios.

As Redes Neurais Artificiais são capazes de solucionar problemas de modelagens sofisticadas, são capazes de modelar funções complexas e computacionais e apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes. Trata-se de uma estrutura que não possuem formas algorítmicas (que não possuem uma seqüência

de instruções), trabalham de uma forma onde os neurônios “nós” recebem vários sinais de entrada e realizam um só sinal de saída.

Desta forma as Redes Neurais Artificiais são representações versáteis de sistemas, pois ao invés de se relacionar os resultados de uma forma estruturada de informações, basta apresentar os dados de entrada que então os resultados são desenvolvidos pela rede gerando através das técnicas de aprendizado os resultados necessários para se fazer uma previsão.

Esta técnica se torna muito mais interessante pelo fato de ser tolerante a ruídos, pois aprende através de exemplos. Para utilização das Redes Neurais Artificiais no desenvolvimento de sistemas fica simplificado por ser necessário apenas que se indique uma base de dados de resultados para que a rede possa sugerir previsões diante de estudos feitos por ela mesma.

Uma das características mais importantes de uma rede neural artificial é a capacidade de aprender com o meio ambiente, adaptando-se às condições do meio. Isto é feito através de um processo iterativo de ajuste de seus pesos, chamado de aprendizado.

Quando a rede aprende uma solução para uma classe generalizada de problemas dizemos que ela aprendeu a tratar destes. Para que uma rede neural possa fazer isto é necessário fornecer a ela um conjunto de valores possíveis de entradas e, no caso de redes com aprendizado supervisionado, as saídas corretas para as soluções propostas.

Aplica-se então um algoritmo que irá ajustar os pesos das arestas internas da rede, para generalizar aquela classe de problema. Estes algoritmos diferenciam-se entre si principalmente pelo modo como os pesos são ajustados, na Figura 4.15 abaixo ilustramos nosso modelo.

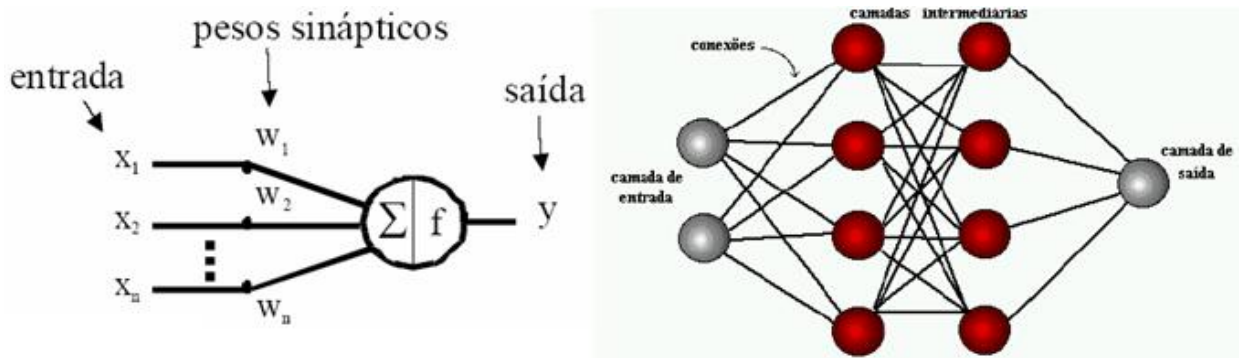


Figura 4-15 - Modelo rede neural e seus elementos (PORTILHO, *et al.* 2006).

Todo o conhecimento de uma rede está nas sinapses, ou seja, nos pesos das arestas. Para que a rede não seja configurada de maneira errada o processo de aprendizado deve ser feito de forma rigorosa. Existem vários paradigmas de aprendizado, dentre os quais se destacam:

- Independência de quem aprende: rede aprende por memorização, contato, exemplos, por analogia, por exploração e também por descoberta;
- Por retroação do mundo: diz respeito à presença ou ausência de realimentação explícita de fora da rede.
- Por Finalidade do Aprendizado:
 - ✓ **Auto-associador:** é apresentada à rede uma coleção de exemplos para que ela memorize. Quando se apresenta um dos elementos da coleção de exemplos, mas de modo errôneo, a rede deve mostrar o exemplo original, funcionando assim como um filtro;
 - ✓ **Hetero-associador:** é uma variação do auto-associador, mas que se memoriza um conjunto de pares. O sistema aprende a reproduzir o segundo elemento do par, mesmo que o primeiro esteja pouco modificado, funcionando desta maneira como um reconhecedor de padrões.

Dentre os principais procedimentos para aprendizado, destacam-se:

- **Aprendizado supervisionado:** aquele no qual um “professor” provê valores de saída para cada padrão de entrada da rede.

Podemos citar como um exemplo o reconhecimento de padrões, os quais devem ser fornecidos por um operador (professor). Os parâmetros de entrada serão: cor, tamanho e textura, e as saídas previstas podem ser somente: peixe bom ou peixe estragado. Podemos visualizar o exemplo pela Figura 4.16.

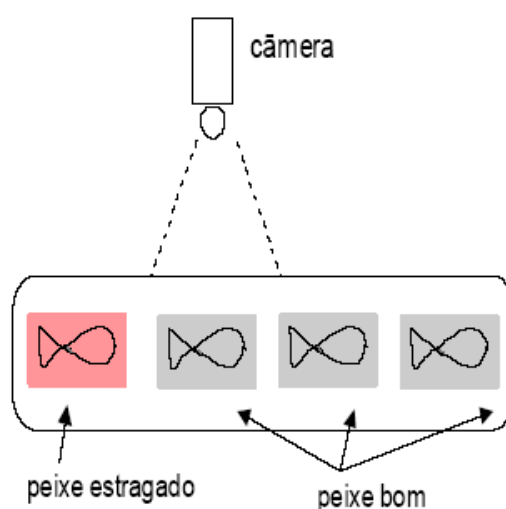


Figura 4-16 – Esquema de controle de padrão industrial (PORTILHO, et al. 2006).

- **Aprendizado semi-supervisionado:** aquele no qual um professor meramente indica quando a resposta da rede a um padrão de treinamento é “bom” ou “ruim”;
- **Aprendizado não-supervisionado:** no qual não há um “professor”, e a rede deve encontrar regularidades nos dados de treinamento por si mesma.

Algo entre 50 e 90% das respostas devem ser fornecidas no processo de aprendizado, afim de que a rede aprenda a resolver o problema e não apenas decore soluções prontas.

Podemos denominar ainda ciclo como sendo uma apresentação de todos os N ($N \in \mathbb{N}$, e $N > 0$) pares (sendo N as entradas e saídas) do conjunto de treinamento no processo de aprendizado.

A correção dos pesos num ciclo pode ser executada de dois modos:

- **Modo Padrão:** A correção dos pesos acontece a cada apresentação à rede de um exemplo do conjunto de treinamento. Cada correção de pesos baseia-se somente no erro do exemplo apresentado naquela iteração. Assim, em cada ciclo ocorrem N ($N \in \mathbb{N}$, e $N > 0$) correções.
- **Modo Batch:** Apenas uma correção é feita por ciclo. Todos os exemplos do conjunto de treinamento são apresentados à rede, seu erro médio é calculado e a partir deste erro fazem-se as correções dos pesos.

Redes neurais artificiais são geralmente utilizadas para fazer reconhecimento estático de padrões, ou seja, para estaticamente mapear entradas complexas em saídas mais simples, como classificação N -ária de padrões de valores.

O modo usual de se treinar uma rede neural para a tarefa a ser executada é através de um procedimento chamado de backpropagation, onde os pesos da rede são modificados em proporção de sua contribuição ao erro observado nas unidades de saída da rede (RUMELHART; HINTON e WILLIAMS, 1986).

Neste contexto, a aprendizagem é um processo pelo qual os parâmetros livres de uma RNA são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. Um conjunto pré definido de soluções para um treinamento de uma Rede Neural é denominada algoritmo de aprendizagem. Não há um algoritmo único de aprendizagem para se projetar uma RNA. De fato, há uma variedade de algoritmos de aprendizagem, cada qual oferecendo vantagens específicas. Um algoritmo de aprendizagem define basicamente a estrutura de treinamento da Rede Neural (HAYKIN, 1999).

Porém, conforme Haykin (1999), uma dificuldade quanto à aplicação prática das RNA decorre quando queremos determinar um modelo de RNA para fazer um estudo específico. Na verdade, a escolha de um modelo de RNA particular é influenciada pela tarefa de aprendizagem que ela deve executar. De maneira geral, seis tarefas de

aprendizagem podem ser executadas utilizando-se RNA: associação de padrões, reconhecimento de padrões, aproximação de funções, controle, filtragem e formação de feixe. Destas, indicam que durante a prática clínica utiliza-se mais freqüentemente a função associativa de padrões. De fato, sabe-se que a associação é uma característica proeminente da memória humana, e todos os modelos de cognição utilizam a associação de uma forma ou de outra como operação fundamental (ANDERSON *et al.*, 1995).

Para Reggia (1993), particularmente, a regra de modelar sistemas para a decisão médica pode ser similar a métodos estatísticos para reconhecimento de padrões. O reconhecimento de padrões, também denominado classificador de padrões, é formalmente conceituado como o processo pelo qual um padrão ou sinal recebido é atribuído a uma classe dentre um número pré-determinado de classes, também denominadas categorias. Uma RNA realiza a identificação de padrões passando inicialmente por uma seção de treinamento, durante a qual se apresenta repetidamente à rede um conjunto de dados de entrada junto com a categoria a qual cada padrão particular pertence. Posteriormente, a rede é capaz de identificar a classe de um novo padrão devido à informação que ela extraiu dos dados de treinamento (HAYKIN, 1999).

Para iniciar um processo de desenvolvimento de uma Rede Neural são realizados dois passos inicialmente, a coleta de dados e a separação de conjuntos. A coleta de dados é feita para que a rede possa usar esses dados para uma aprendizagem inicial, logo em seguida a separação de conjunto de dados é feito para se obter as relações entre os dados projetados pela rede. O terceiro passo é a configuração da rede, onde devemos levar em consideração três subdivisões importantes nesse processo, que são:

- Seleção do paradigma neural apropriado à aplicação.
- Determinação da topologia da rede a ser utilizado - o número de camadas, o número de unidades em cada camada, etc.
- Determinação de parâmetros do algoritmo de treinamento e funções de ativação. Este passo tem um grande impacto no desempenho do sistema resultante.

No quarto passo que é a parte do treinamento da rede, entra em execução o algoritmo escolhido onde serão ajustados os pesos das conexões. É bem relevante que se julgue os fatores como inicialização da rede e o tempo necessário para execução do treinamento. Quanto ao tempo de treinamento vários fatores podem levar ao tempo de duração, porém, deve-se definir um critério de parada, pois o algoritmo de backpropagation a sua parada não é bem definida. Então se leva em conta o número de iterações do ciclo e sua taxa de erro. O treinamento deve ser interrompido quando houver uma boa taxa de generalização e uma pequena taxa de erro.

O quinto passo é o teste da rede. Durante esta fase o conjunto de teste é utilizado para determinar o desempenho da rede com dados que não foram previamente utilizados. O desempenho da rede, medida nesta fase, é uma boa indicação de seu desempenho real.

Finalmente, depois da fase de teste da rede ela já se encontra pronta para ser implementada em um sistema que possua facilidade para o usuário, permitindo seu entendimento para que futuramente com a aquisição de vários dados seja talvez a necessidade de um novo treinamento para ajustamento da rede.

4.7. ALGORITMOS GENÉTICOS

4.7.1. Introdução

Em meados dos anos 50 e 60, começou-se a desenvolver simulações computacionais dos sistemas genéticos. Neste sentido pesquisadores buscavam através da natureza inspiração para certas soluções. Observando-a conseguiram distinguir sua capacidade de resolver problemas, como por exemplo, a maneira como ela resolve o fato da sobrevivência das espécies, onde as adaptações dos seres vivos se davam através do processo de natureza estatística como seleção, cruzamento e mutação (FOGEL, 1995).

No entanto, John Holland (*apud* Azevedo, 1975) que iniciou as pesquisas sobre este tema, gradualmente foi implementando sua idéia e em 1975 publicou seu livro *Adaptation in Natural and Artificial Systems*. Atualmente é considerado a bíblia dos Algoritmos Genéticos, suas teorias foram inspiradas nos estudos de Darwin e Mendel, onde Darwin citava sobre a evolução das espécies na qual descreve que um indivíduo sofre influencia do meio onde vive, desta forma os mais aptos terão maior probabilidade de sobrevivência em tal ambiente. (Darwin, 1981 *apud* Azevedo, 1999). Por outro lado, Mendel afirma que o material genético dos pais é passado aos seus descendentes, gerando assim novos indivíduos mais adaptados com o ambiente (DUNN, 1950 *apud* AZEVEDO, 1999).

Daí por diante estes algoritmos vêm sendo aplicados em diversas áreas que necessitem da utilização de sistemas de otimização, aprendizagem de máquina e principalmente na bioinformática como, por exemplo, na descoberta do mapeamento do DNA, além de serem considerados bastante eficazes para otimização na busca por resultados de ordem numérica, otimização de funções, aprendizado de máquina (FORGEL, 1995), na classificação de dados e podendo ser utilizada para recuperação ou para efetuar análise de dados (FURTADO, 1998).

Alguns fatores fizeram com que os Algoritmos Genéticos se destacassem bastante como uma técnica de sucesso (SILVA, 1995 *apud* AZEVEDO, 1999), (TANOMARU, 1995 *apud* AZEVEDO, 1999):

- Simplicidade de operações;

- Facilidade de implementação;
- Eficácia na busca da região onde, provavelmente, encontra-se o máximo global;
- Aplicável em situações onde não se conhece o modelo matemático ou este é impreciso e também em funções lineares e não lineares;
- AGs trabalham com uma codificação do conjunto de parâmetros e não com os próprios parâmetros;
- AGs trabalham com uma população e não com um único ponto;
- AGs utilizam informações de custo ou recompensa e não derivadas ou outro conhecimento auxiliar;
- AGs utilizam regras de transição probabilísticas e não determinísticas.

Os Algoritmos Genéticos são métodos bastante eficientes para utilização em problemas que necessitam de soluções ótimas ou quase ótimas, devido a não impor muitas limitações para suas soluções, diferindo neste aspecto dos métodos de busca tradicionais. O Algoritmo Genético utiliza-se de uma estratégia eficiente que torna seus resultados ainda mais vantajosos, que é a técnica de Gerar-e-Testar, além de serem baseados nas evoluções biológicas que tornam capazes de identificar e explorar fatores ambientais e transformá-los em soluções ótimas ou quase ótimas.

Significa que quanto melhor um indivíduo se adaptar ao seu meio ambiente, maior será sua chance de sobreviver e gerar descendentes: este é o conceito básico da evolução genética biológica. Portanto, Algoritmos Genéticos são desenvolvidos através de uma simulação do fenômeno evolutivo das espécies, conforme descrito por Darwin.

4.7.2. Fundamentos Biológicos

Todos os seres vivos são constituídos de células, onde cada célula possui uma cadeia de cromossomos, onde são constituídos de moléculas de DNA. Os cromossomos possuem genes, tais genes são uma representação da função local do DNA e a função principal de cada gene é codificar algum tipo de proteína para ser absorvida pelo organismo. Com isso podemos dizer que cada gene tem função de determinar alguma característica de cada ser, como por exemplo, cor dos olhos, tipo de cabelo e etc. O

conjunto de genes relacionado com alguma característica é chamado de alelos. Cada gene tem sua posição dentro do cromossomo, essa posição é chamada de local.

Um conjunto completo de todo o material genético (todos os cromossomos) é chamado de genoma. Um conjunto particular de genes de um genoma é chamado de genótipo. O genótipo mais todo o desenvolvimento do organismo após o seu nascimento é a base do fenótipo, que são todas as características físicas e mentais que o organismo adquire durante toda a sua vida, tais como: cor de olhos, inteligência, etc.

A seleção natural depende muito das condições ambientais em que os seres vivem. Diante da adaptação que os seres vivem, quem melhor se adapta adquire melhores condições para sobreviver e conseqüentemente ajudando na reprodução. Já os seres menos adaptados podem vir a morrer mesmo antes de conseguirem a sua reprodução. A geração que se formará reproduzida pelos seres mais evoluídos carregam as características de adaptação, assim, caso o meio em que vivem não sofra alterações os seres mais adaptados terão um maior número de descendentes através de várias iterações, podendo os organismos virem a desenvolver características adaptativas bastante complexas.

A teoria da seleção natural não é apenas aplicada a organismos biológicos, mas a qualquer organismo que se reproduza de modo a envolver tanto a hereditariedade como a variação (MICHALEWICKS, 1996).

Desta forma, pode ocorrer seleção natural no reino não biológico. Vários pesquisadores buscam inspiração na natureza para desenvolver técnicas de busca para uma série de problemas.

Na natureza o processo de seleção natural demonstra que seres mais preparados (aptos) competem pelos recursos naturais, tendo assim uma maior probabilidade de sobreviver, conseqüentemente, disseminar o seu código genético (SRINIVAS & PATNAIK, 1994).

Com o passar das gerações, através de inúmeros cruzamentos e mutações que ocorrem com as espécies, estas tendem a estar cada vez mais adaptadas ao meio ambiente em que vivem.

Emprega-se no AG a mesma terminologia e os mesmos princípios da teoria evolutiva e da genética. O AG relacionado com a seleção natural pode ser expresso como: (DARWIN, 2004)

- SE há organismo que se reproduzem;
- SE os descendentes herdam as características de seus genitores;
- SE há variação nas características;
- SE o ambiente não suporta todos os indivíduos de uma população em crescimento;
- ENTÃO os indivíduos que apresentarem menor adaptação (determinadas pelo ambiente) morrerão;
- ENTÃO os indivíduos que apresentarem maior grau de adaptação (determinadas pelo ambiente) prosperarão.
- Como resultado desse processo tem-se a evolução das espécies.

4.7.3. Características globais dos algoritmos genéticos

Os Algoritmos Genéticos são algoritmos que possuem como principal tarefa a otimização global, tendo como fonte de inspiração os mecanismos de seleção natural e da genética. Eles trabalham de forma paralela e estruturada, porém, aleatória, onde visam o reforço dos pontos da “alta aptidão”, pontos nos quais a função deve ser minimizada ou maximizada, encontrando sempre soluções para problemas observando sempre estes fatores.

Por serem de forma aleatória não significa que sua forma não seja bem direcionada, pois trabalham com informação passadas, onde seus próximos passos sempre envolvem uma exploração de fatores capazes de encontrar sempre os melhores resultados, isto ocorre sempre a cada processo de iteração, onde nesse caso a iteração é abordada como geração.

A cada interação, as técnicas para a seleção e reprodução de uma população de indivíduos podem variar, isso depende do problema a ser solucionado e de recursos para tal aplicação. Utilizando o método de Seleção, se torna capaz de classificar os indivíduos que irão se reproduzir, gerando seus descendentes de acordo com a capacidade de cada indivíduos de se reproduzir, levando em conta que apenas os mais aptos conseguem se reproduzir, assim consegue-se calcular estatisticamente a capacidade de reprodução de cada indivíduos. Indivíduos com maior aptidão têm maiores chances de se reproduzir.

O principal fator relevante para a escolha da utilização do Algoritmo Genético para uma solução é sua forma de representação, que utiliza de fatores históricos para dar seqüência em suas tarefas, além da utilização de vetores finitos em um alfabeto finito.

De maneira geral, cada individuo recebe uma denotação genotípica formando-se então um vetor binário, onde cada valor binário é representado por presença = 1 ou ausência = 0, representado de acordo com seu genótipo. Os elementos podem ser combinados formando características reais do individuo. Essa forma de representação não é característica específica do problema, a partir do momento que é caracterizada esta simbolização binária ela poderá ser facilmente utilizada em diferentes classes do problema.

O critério de seleção utilizado pelos Algoritmos Genéticos é que irá depois de várias gerações, originar indivíduos mais aptos, esse é um dos fundamentos principais dos Algoritmos Genéticos. Os métodos de seleção procuram selecionar sempre os indivíduos com maiores notas de aptidão, porém, não exclusivamente, para sempre manter a diversidade da população.

Uma maneira adequada para escolha de indivíduos é o Método da Roleta, onde indivíduos são sorteados aleatoriamente, e em seguida selecionados para fazerem parte da próxima geração.

Desta maneira, os indivíduos são representados de acordo com sua aptidão, ou seja, se um individuo tem maior aptidão o tamanho dele no espaço da roleta será maior, e se a aptidão dele for pequena, seu espaço será relativamente menor. Finalmente, a roleta é girada várias vezes, isso varia de acordo com o tamanho da população e então, são selecionados para a próxima geração os indivíduos sorteados na roleta.

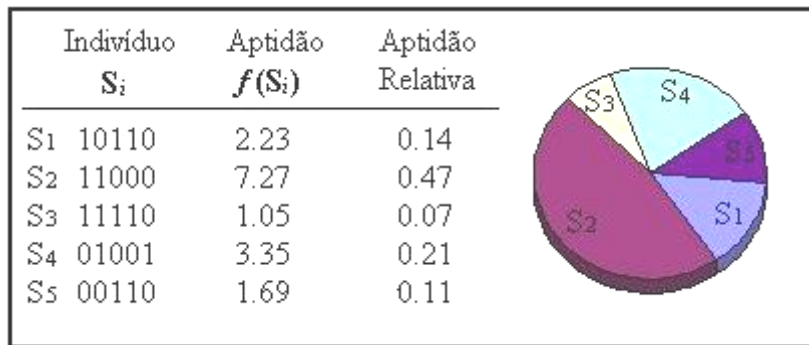


Figura 4-17 – Indivíduos de uma população e a sua correspondente roleta de seleção.

Para que uma população consiga gerar suas descendências é necessário que haja algumas operações, até mesmo para que melhore a aptidão desta população com o tempo. Essas operações são o cruzamento e a mutação, que são responsáveis por gerar uma nova população, mas que possua as características da geração anterior, carregando as características de seus pais, diversificando, mantendo as características de seus pais e trazendo adaptações existentes da geração anterior. Para assegurar que os melhores indivíduos não desapareçam nas próximas gerações devido às operações genéticas, eles podem ser automaticamente adicionados à próxima geração, através da reprodução elitista.

Esse ciclo de gerações é repetido por um determinado numero de vezes, segue abaixo um exemplo de Algoritmo Genético, onde podemos ver que durante o processo, os indivíduos e alguns dados estatísticos são coletados e armazenados para avaliação.

Procedimento AG

{ t = 0;

inicia_população (P, t)

avaliação (P, t);

repita até (t = d)

{ t = t +1;

seleção_dos_pais (P,t);

recombinação (P, t);

mutação (P, t);

avaliação (P, t);

sobrevivem (P, t)

}

}

onde:

t - tempo atual;

d - tempo determinado para finalizar o algoritmo;

P - população

Estes tipos de algoritmos, apesar de serem aparentemente muito simples, realizam inúmeros tipos de tarefas altamente complexas e se tornam muito poderosos, além de não serem limitados sobre o espaço de busca.

4.7.4. Operadores Genéticos

O objetivo principal dos operadores genéticos é gerar, observar e transformar a população utilizando-se de sucessivas e inúmeras gerações, estendendo a busca até chegar a um resultado que satisfaça a solução. Os operadores genéticos têm um papel principal e são altamente necessários por manterem a diversidade e as características de adaptação adquirida pelas gerações anteriores.

Para que haja sempre uma preservação da diversidade genética dentro da população se torna sempre necessário a utilização dos operadores genéticos, onde ele tem o papel de alterar de maneira responsável os componentes de uma estrutura escolhida. Como demonstrado na Figura 4.18 ele proporciona a possibilidade de introduzir novos indivíduos dentro da população, assim a mutação da toda segurança que as buscas dentro de um espaço nunca serão zero.

Antes da Mutação:	1 1 (1) 0 0
Depois da Mutação:	1 1 0 0 0

Figura 4-18 – Exemplo de mutação

O operador responsável pela recombinação das características dos pais durante a reprodução é o cruzamento, dando a possibilidade assim que as próximas gerações herdem suas características. O cruzamento é considerado o operador genético mais importante, é

por isso que se torna mais utilizado, atuando sempre diante da probabilidade da taxa de crossover, que deve ser maior que a taxa de mutação. O cruzamento pode ser utilizado de variadas maneiras, as mais importantes são:

- **Um-ponto:** um ponto de cruzamento é escolhido e a partir deste ponto as informações genéticas dos pais serão trocadas. As informações anteriores a este ponto em um dos pais são ligadas às informações posteriores a este ponto no outro pai, como é mostrado no exemplo da Figura 4.19:
- **Multi-pontos:** é uma generalização desta idéia de troca de material genético através de pontos, onde muitos pontos de cruzamento podem ser utilizados.
- **Uniforme:** não utiliza pontos de cruzamento, mas determina, através de um parâmetro global, qual a probabilidade de cada variável ser trocada entre os pais.

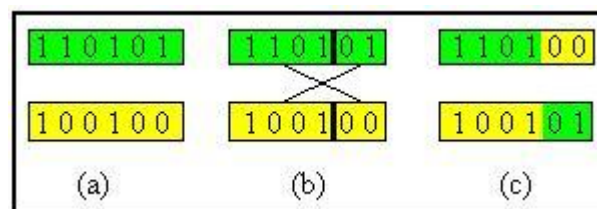


Figura 4-19 – Um exemplo de crossover de um ponto. (a) dois indivíduos são escolhidos. (b) um ponto de crossover é escolhido. (c) são recombinações das características, gerando dois novos indivíduos.

4.7.5. Parâmetros Genéticos

É de grande valia observar como os parâmetros podem influenciar nos comportamentos das atividades dos Algoritmos Genéticos, desta forma os parâmetros tornam-se necessários para estabelecer as ações que serão necessárias para as soluções dos problemas de acordo com os recursos necessários.

O tamanho da população pode atrapalhar o desempenho global e a eficiência do AG, pois com uma população reduzida os resultados podem não ser tão favoráveis pelo valor pequeno de diversidade que possa existir, diminuindo o espaço de ação.

Diferentemente de uma população relativamente grande, onde o AG é capaz de obter várias opções de variabilidade, evitando convergências prematuras. Porém, em uma grande população exigem-se maiores recursos computacionais e tempo de processamento.

A taxa de cruzamento são parâmetros importantes, quanto maior esta taxa mais rapidamente serão criadas novas estruturas para a atual população, porém se for muito alta essa taxa pode atrapalhar na performance da população, pois pode acontecer da população atual ser totalmente substituída, além dos indivíduos de boa aptidão serem substituídos em algum cruzamento. Já com um valor baixo o algoritmo se torna muito lento, pois acaba não gerando novos indivíduos não tendo um bom crescimento.

As taxas de mutação podem ser baixas ou altas, utilizando uma baixa taxa inibe que uma dada posição fique estagnada em um único valor, podendo assim ter a possibilidade de chegar em qualquer ponto do espaço de busca. Já com uma taxa muito alta a busca se torna essencialmente muito alta.

Os intervalos de gerações controlam as porcentagens de populações que serão substituídas durante a próxima geração. Se esse intervalo de geração for muito alto a maior parte da população será substituída, ocasionando assim a perda de aptidão, pois, indivíduos com uma boa aptidão serão substituídos nesse momento. Já com um valor baixo o algoritmo se torna lento.

4.7.6. Pesquisas e Aplicações sobre AG

A cada dia que se passa os Algoritmos Genéticos vem sendo cada vez mais utilizados. Sua forma de aplicação e seus resultados vão se tornando a cada dia essenciais em sistemas com problemas cada vez mais complexos. Podendo assim serem vistos em sistemas como aprendizado de máquina além de serem observados em sistemas simples também. O mais interessante hoje são sua utilização em sistemas artísticos como a pintura e a música.

Suas vantagens e suas formas de aplicação se tornam interessantes pelo fato de usar de uma forma paralela, percorrendo todo o espaço da solução, reunindo informações e

obedecendo a parâmetros que acabam tornando seus resultados cada vez mais sensíveis a uma boa solução, pois não se prendem apenas a uma local como nos outros métodos.

Os Algoritmos Genéticos possuem uma incrível vantagem por serem de fácil implantação, uma vez implantados a sua parte básica, é necessário apenas descrever os novos objetos a serem utilizados para resolverem os problemas. Na codificação é necessário apenas mudar a função de adequação para enfim dar fim a atualização. Porém em alguns sistemas a adequação pode ser extremamente difícil.

O tempo de processamento dos Algoritmos Genéticos geralmente é razoavelmente mais lento que os outros métodos utilizados, mas se tratando dos resultados e do nível de certeza dos mesmos, o tempo gasto está em um intervalo aceitável principalmente com o desempenho dos computadores que existem hoje.

Portanto, se torna evidente a eficácia da utilização desta técnica bastante satisfatória, onde podem-se exemplificar abaixo, alguns problemas resolvidos com os Algoritmos Genéticos:

- Sistemas Dinâmicos não Lineares - predição, análise de dados;
- Projeto de Redes Neurais, ambos: determinação da arquitetura e dos pesos;
- Trajetória de Robôs;
- Desenvolvimento de programas em LISP (programação genética);
- Planejamento Estratégico;
- Determinação da Forma de Moléculas de Proteínas;
- Determinação de Rotas e Sequenciamento de Tarefas;
- Determinação de Funções para criação de imagens.

5. METODOLOGIA

Neste capítulo descrevem-se os métodos utilizados no sistema de detecção dos sinais nos dados eletroencefalográficos. Ou seja, apresenta-se o sistema que foi proposto e toda a metodologia para o seu desenvolvimento.

5.1. OBTENÇÃO DOS DADOS DE EEG

A obtenção dos dados de exames eletroencefalográficos foi definida pela separação entre os dados de resultados positivos e de resultados negativos. A partir daí realizou-se a simbolização das séries temporais para cada sinal a fim de avaliar os dados gerados para que seja montada uma base de dados com as características principais e eminentes de cada tipo de exame: positivo e negativo.

A separação dos eventos epileptiformes dos sinais normais foi realizada pelo próprio *software* desenvolvido, a partir deste momento a base de dados com duas características distintas encontraram-se aptas a dar seqüência aos demais processos descritos neste trabalho.

A parte selecionada será a parte do sinal a qual será salva na base de dados, notam-se claramente as características exploradas, após isso salvamos estes sinais para aplicarmos à simbolização, onde através da simbolização das séries temporais será possível obter características claras que diferenciem ambos os sinais.

5.2. SIMBOLIZAÇÃO DAS SÉRIES

Para análise e caracterização dos sinais advindos dos exames de EEG, utilizou-se o método de simbolização de séries temporais experimentais descritos por (DAW, 2003), o método consiste em transformar a série temporal em uma seqüência de n símbolos, baseados em um alfabeto. Por exemplo, utilizando um alfabeto binário ($n=2$), a cada medida disponível nos dados são atribuídos os símbolos “0” ou “1”, segundo o valor da medida esteja abaixo ou acima de um dado limiar, respectivamente, como mostrado na

Figura 5.1, para o caso de uma série temporal periódica (senoidal). Em seguida, uma seqüência de símbolos de comprimento m é escolhida, e as freqüências relativas de todas as possíveis seqüências de símbolos de dimensão m são calculadas para a seqüência completa obtida da discretização dos sinais experimentais (ver Tabela 5.1).

Representando o conjunto de comprimento m , encontraremos seu equivalente decimal, com isto, pode-se representar a distribuição estatística das seqüências simbólicas como um histograma, mostrado na Figura 5.2, no qual o eixo horizontal corresponde às seqüências simbólicas e o eixo vertical a freqüência relativa das seqüências.

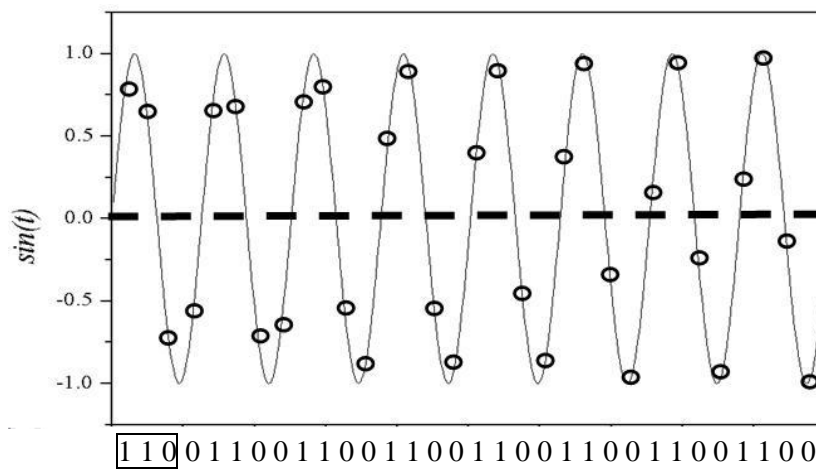


Figura 5-1 – Simbolização de série temporal.

Para dados puramente aleatórios, todas as seqüências simbólicas de comprimento m são equiprováveis. Portanto, o distanciamento deste comportamento indica a evidência de estrutura determinística nos dados, podendo ser epiléptica ou não.

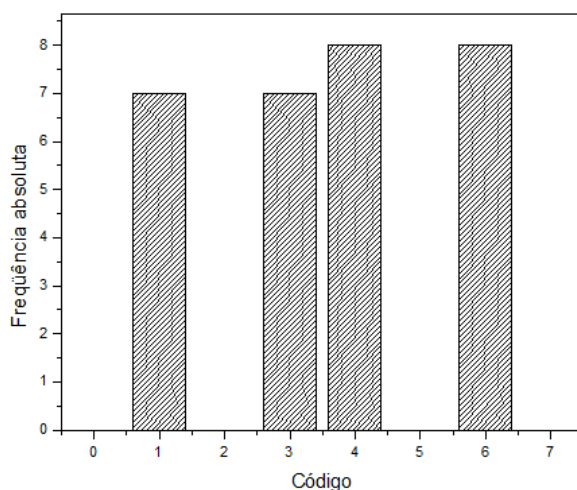


Figura 5-2 – Distribuição de frequência absoluta para uma série temporal periódica.

Seqüência	Código	Frequência
000	0	0
001	1	7
010	2	0
011	3	7
100	4	8
101	5	0
110	6	8
111	7	0

Tabela 5.1 – Frequência das seqüências obtidas na Figura 5-2.

No contexto da simbolização existem quatro parâmetros os quais são de suma importância para que a simbolização possa ser executada em uma ótima performance. Estes parâmetros são: *Tamanho do Alfabeto (TA)*, *Tamanho da Palavra (TP)*, *Tamanho da Janela (TJ)* e *Passo da Janela (PJ)*. Onde o *Tamanho do Alfabeto* define quantos símbolos serão utilizados para simbolização das séries, o *Tamanho da Palavra* implica a quantidade de símbolos que formarão a palavra, o *Tamanho da janela* expressa a seção do sinal que será simbolizada e o *Passo da Janela* é intervalo de tempo entre uma posição da janela temporal e a posição subsequente durante a simbolização do sinal.

Dependendo dos valores atribuídos aos parâmetros pode ocorrer de não ser possível uma distinção, no contexto da simbolização, entre os sinais com epilepsia e sem epilepsia. Para resolver esse tipo de problema, uma ferramenta de otimização é implementada, neste caso os Algoritmos Genéticos irão otimizar os parâmetros necessários para a simbolização das séries. O objetivo é encontrar os parâmetros que proporcionem a melhor distinção entre os sinais com eventos epiléticos e não epiléticos.

No algoritmo de otimização, implementou-se a Entropia de Shannon, uma medida estatística capaz de encontrar graus de incerteza entre os valores utilizados, assim, quanto menor o valor do índice de Shannon menor o grau de incerteza, portanto, comparando os dois tipos de sinais quanto maior a diferença do grau de incerteza entre eles, melhor é a diferenciação entre os sinais, conseqüentemente, mais adequados são os parâmetros utilizados para a simbolização das series temporais naquele instante.

$$H_s(L) = -\frac{1}{\log N_{obs}} \sum_i p_{i,L} \log_{p_{i,L}} \quad (1)$$

Onde N_{obs} é o total de número de colunas do histograma com valor diferente de zero e $p_{i,L}$ é a frequência relativa dos valores de cada coluna do histograma.

Para cada geração de indivíduos criada pelo Algoritmo Genético será calculada a sua entropia onde encontra-se para cada um sua aptidão, assim somaremos as aptidões de todos os indivíduos de certa geração comprovadamente epiléticos e outra não epiléticos, calculando em seguida a diferença entre eles.

$$Dif = \left| \int i_n - \int i_{ep} \right| \quad (2)$$

Após esse processo o algoritmo genético irá encontrar a população com a maior aptidão, isso significa que a população possui uma maior diferença entre epiléticos e não epiléticos. Através desta otimização o Algoritmo Genético nos informará os valores dos parâmetros utilizados para simbolização a qual obteve uma melhor diferença entre os sinais.

5.3. CLASSIFICAÇÃO DOS SINAIS

Para classificação e detecção automática da presença de eventos epileptiformes nos sinais de exames de eletroencefalografia, utilizamos as Redes Neurais Artificiais (RNA), um dos métodos mais utilizados nas literaturas para o tipo de finalidade deste trabalho.

MÉTODOS PARA DETECÇÃO	
DESCRIÇÃO	USABILIDADE (%)
Redes Neurais	61,5%
Threshold	30,7%
Regras Heurísticas	5,7%
Medidas de Distância	1,9%

Tabela 5.2 – Tabela da usabilidade dos algoritmos para detecção de crises (Extraído de PARREIRA, 2006).

É importante salientar que esta metodologia pode ser aplicada juntamente com outros métodos, podendo aumentar a capacidade de detecção. Em Redes Neurais, pode-se encontrar diversas arquiteturas de RNAs que são aplicadas nas diversas detecções. Na Tabela 5.3 ilustra as diversas metodologias utilizadas na literatura em geral.

MÉTODOS PARA DETECÇÃO COM REDES NEURAIAS	
DESCRIÇÃO	USABILIDADE (%)
<i>Backpropagation</i>	23,0%
MLP	15,3%
PNN	5,7%
RBF	3,8%
<i>Backpropagation time delay</i>	3,8%
ART	3,8%
SOM	1,9%
TDNN	1,9%

Tabela 5.3 – Redes neurais mais utilizadas – (Extraído de PARREIRA, 2006).

Segundo Parreira (2006), observa-se na Tabela 5.3, que o algoritmo de classificação da Rede Neural mais utilizado foi a Backpropagation com 23,0%, logo após a MLP com resultado de 15,3% e abaixo a PNN com 5,7%. As outras Redes Neurais, apesar de alguns

pesquisadores terem utilizado e conseguido êxito na classificação dos eventos, tiveram baixa usabilidade para este tipo de pesquisa.

Apesar do método Backpropagation ser o mais utilizado, neste trabalho utilizou-se o método Perceptron Multicamada (MLP). De acordo com estatísticas, seu uso se encontra em segundo lugar, porém, consegue realizar a classificação de forma mais simples de acordo com a técnica proposta.

O Perceptron Multicamada pode ser adicionado a diversas camadas ocultas. Estas camadas têm como objetivo, melhorar o desempenho da rede, aumentando a possibilidade de divisão de espaço de entrada de maneira não-linear. Com a possibilidade de ampliação de mais camadas, pode-se alcançar um número maior de pesos ou sinapses. Fazendo um paralelo com o parâmetro biológico, dentro do tecido nervoso, já se prova que com o passar do tempo, os neurônios aumentam as conexões sinápticas entre si para efetuarem melhor seus mecanismos de troca de íons. Esse aumento de pesos ou sinapses faz com que as trocas de informações e os cálculos sejam maiores, atingindo um resultado bem próximo do esperado (MEDEIROS, 2006).

No MLP existe entre a camada de entrada e a camada de saída uma ou mais camadas ocultas. A camada oculta leva esse nome devido a ela não possuir acesso direto à saída da rede MLP, onde os erros de aproximação são calculados.

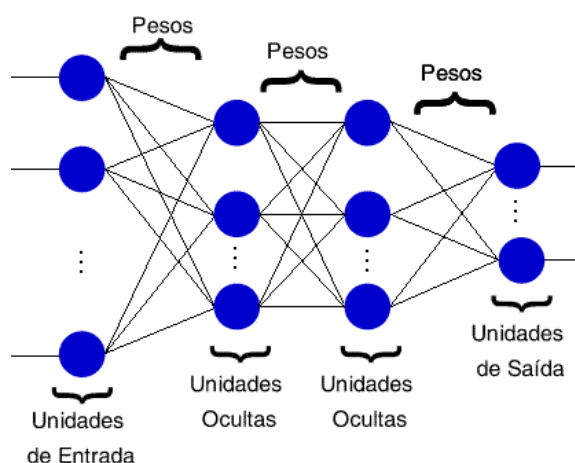


Figura 5-3 – Rede Neural Perceptron Multicamadas.

Na rede MLP existem duas fases distintas denominadas de treinamento/aprendizagem e teste/prova. Estas fases são bastante utilizadas e essenciais para o funcionamento adequado da rede. Para que uma RNA possa fornecer resultados convenientes, é necessário que seja treinada, onde seus pesos são ajustados de forma que ela se adapte aos diferentes estímulos de entrada. Durante esta fase de treinamento, ocorre o seu aprendizado.

Existem vários processos de aprendizado, neste trabalho utiliza-se o *Aprendizado Supervisionado*, onde informamos à RNA uma referência do objetivo a ser alcançado. Dentro do *Aprendizado Supervisionado* executa-se um processo conhecido como *Aprendizado por Correlação de Erro*, onde o erro de uma RNA (Equação 3) pode ser calculado como a diferença entre a saída real gerada pela rede e a saída desejada, fornecida através de um treinamento supervisionado, por exemplo:

$$e_k = d_k - y_k \quad (3)$$

Onde para um estímulo k , e é o sinal de erro, d é a saída desejada apresentada durante o treinamento e y é a saída real da rede após a apresentação do estímulo de entrada. Durante o *Aprendizado Supervisionado*, os erros vão sendo calculados sucessivamente até que se aproximem dos valores satisfatórios. O processo de *Aprendizado por Correção de Erros* utiliza algoritmos para caminhar sobre a curva de erros, com o intuito de alcançar o menor valor de erro possível, o mínimo global (Figura 5.4). Muitas vezes, o algoritmo não alcança este mínimo global, atingindo o que chamamos de mínimo local. Caso este erro alcançado seja desfavorável, é necessário recomeçar processo de aprendizado.

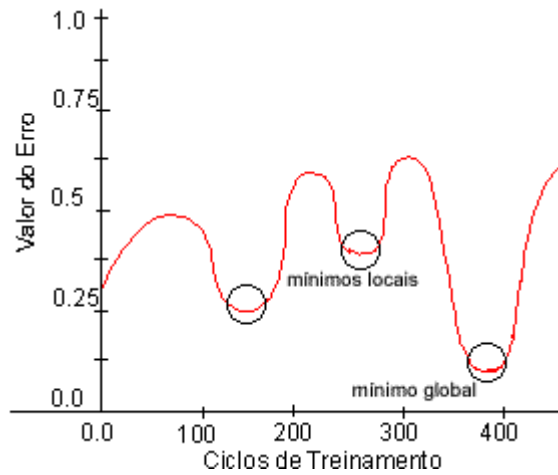


Figura 5-4 – Gráfico de uma possível superfície de erro mostrando os mínimos locais e o mínimo global.

Para a correção do erro, os pesos da rede devem ser ajustados, de forma a aproximar a saída real à desejada. De acordo com a Regra Delta de Aprendizado (Equação 4), tal ajuste dependerá do próprio erro calculado; do valor do estímulo de entrada que é "transmitido" pelo peso a ser ajustado; e também da taxa de aprendizado, a qual relaciona-se à cautela com que a curva de erros é percorrida. Para um dado estímulo k , no passo de treinamento n :

$$\Delta w_{ij} = \eta e_k(n) x_j(n) \quad (4)$$

Onde, $\Delta w(n)$ se refere ao valor de ajuste a ser acrescentado ao peso w_{ij} , η é a taxa de aprendizado, $e(n)$ é valor do erro e $x_j(n)$ é o valor do estímulo. Assim o valor atualizado do peso será:

$$w(n+1) = w(n) + \Delta w_{kj}(n) \quad (5)$$

Portanto, podemos utilizar a Regra Delta (Equação 6) para corrigir os valores dos pesos, minimizando a função de erro \mathcal{E} , também conhecida como "função de custo":

$$\mathcal{E}(n) = \frac{1}{2} e^2(n) \quad (6)$$

Onde, $e(n)$ representa o erro da rede no passo n do treinamento e $\mathcal{E}(n)$ o valor da função de custo no passo n do treinamento. Estas técnicas acima citadas representam o treinamento de nossa MLP, onde é preparada para que se torne capaz de se aproximar ao máximo de nosso objetivo descrito, mas antes destes critérios descreve-se os perfis que desejados que a RNA classifique:

PERFIL	DESCRIÇÃO
Epiléptico	Sinal com a presença de eventos epilépticos
Não Epiléptico	Sinal sem a presença de eventos epilépticos

Tabela 5.4 – Tabela dos perfis dos sinais.

5.4. ÍNDICES DE DESEMPENHO

As avaliações podem ser representadas de acordo com as seguintes medidas especificadas abaixo:

- **VERDADEIROS POSITIVOS (VP):** eventos que foram identificados pelo sistema como eventos epilépticos e confirmados pela marcação de um especialista.
- **FALSOS NEGATIVOS (FN):** eventos que foram identificados pelo sistema como não sendo eventos epilépticos, porém, os especialistas marcaram como sendo epilépticos.
- **VERDADEIROS NEGATIVOS (VN):** eventos que foram identificados pelo sistema como não sendo eventos epilépticos e confirmados pelos especialistas.
- **FALSOS POSITIVOS (FP):** eventos que foram identificados pelo sistema como sendo eventos epilépticos, mas não sendo confirmados pelos especialistas.

Existem outras formas de medidas de desempenho para expressar a relação entre eventos epilépticos e não epilépticos (SUBASIA, *et al.*, 2005), tais como: sensibilidade (Equação 7) (relação verdadeira positiva) que avaliam a precisão do sistema em detectar eventos epilépticos e a especificidade (Equação 8) (relação verdadeira negativa) avaliam a precisão do sistema em detectar eventos não epilépticos.

$$\textit{Sensibilidade} = \frac{VP}{VP + FN} \times 100[\%] \quad (7)$$

$$\textit{Especificidade} = \frac{VN}{VN + FP} \times 100[\%] \quad (8)$$

A seletividade positiva (Equação 9), ou valor preditivo positivo (VPP), é a capacidade do sistema em indicar corretamente os sinais epilépticos. A seletividade negativa (Equação 10), ou valor preditivo negativo (VPN), é a capacidade do sistema em indicar corretamente os sinais não epilépticos.

$$\textit{Seletividade positiva} = \textit{VPP} = \frac{VP}{VP + FP} \times 100[\%] \quad (9)$$

$$\textit{Seletividade negativa} = \textit{VPN} = \frac{VN}{VN + FN} \times 100[\%] \quad (10)$$

Para estabelecer capacidade do sistema em reconhecer os sinais epilépticos e os sinais não epilépticos é aplicado o quantificador (Equação 11) para reconhecer o sucesso do sistema, esse quantificador é denominado Precisão (SUBASIA, *et al.*, 2005).

$$\textit{Precisão} = \frac{\textit{Sensibilidade} + \textit{Especificidade}}{2} \quad (11)$$

Certamente existem dentro da literatura outras formas de quantificar os resultados, porém estas descritas são as utilizadas neste trabalho.

5.5. MODELAGEM DO SOFTWARE

A metodologia empregada na modelagem do sistema computacional foi baseada no paradigma de orientação a objetos, e sua implementação foi feita em linguagem de programação DELPHI 7, pois se trata de uma linguagem que produz *softwares* com alta qualidade, trazendo ao usuário facilidades de manipulação e uma grande variedade de componentes que facilitam o desenvolvimento necessário para este trabalho.

A principal vantagem da orientação a objetos é que a estrutura do programa torna-se mais simples, devido ao aumento de complexidade da estrutura de dados chamada objeto. Isto facilita significativamente a reutilização, a manutenção e a atualização do código desenvolvido, pois encapsula os trechos mais complexos em objetos que podem ser facilmente manipulados.

6. RESULTADOS

Neste capítulo são apresentados os resultados obtidos de acordo com os métodos aplicados neste sistema, junto com suas respectivas observações e discussão, mostrando comparações entre os diversos sinais analisados e suas respectivas análises estatísticas mediante o estudo de caso realizado.

6.1. APRESENTAÇÃO DO SOFTWARE

A Figura 6.1 demonstra as funcionalidades do sistema e em seguida as telas que foram desenvolvidas.

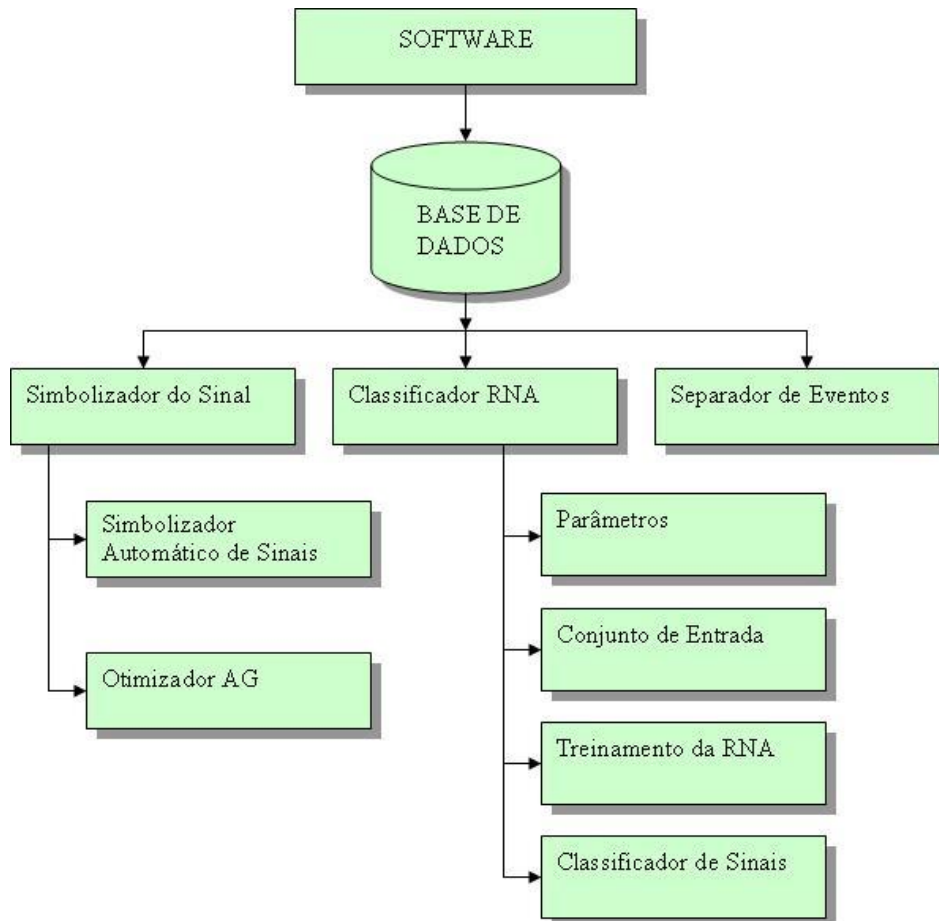


Figura 6-1 – Diagrama de blocos do sistema, mostrando suas possibilidades de navegação.

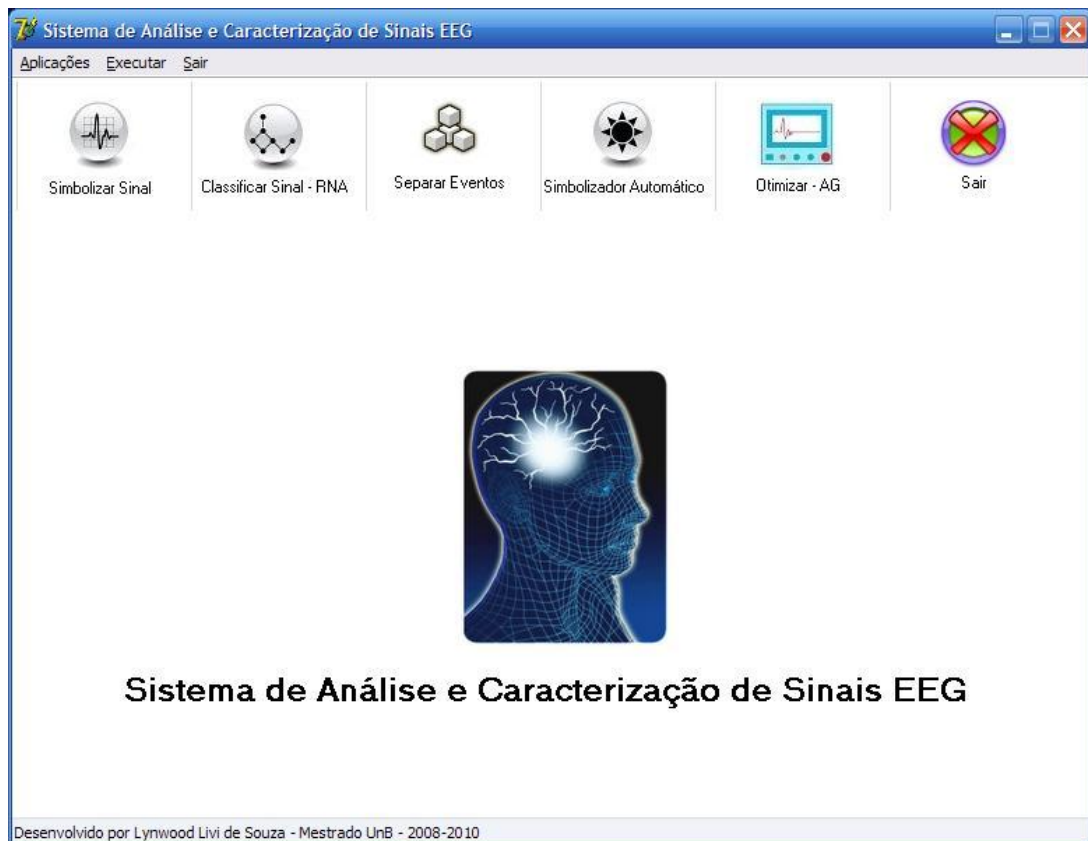


Figura 6-2 – Tela principal do Sistema.

A tela principal oferece opções ao usuário de acessar todas as áreas do sistema.

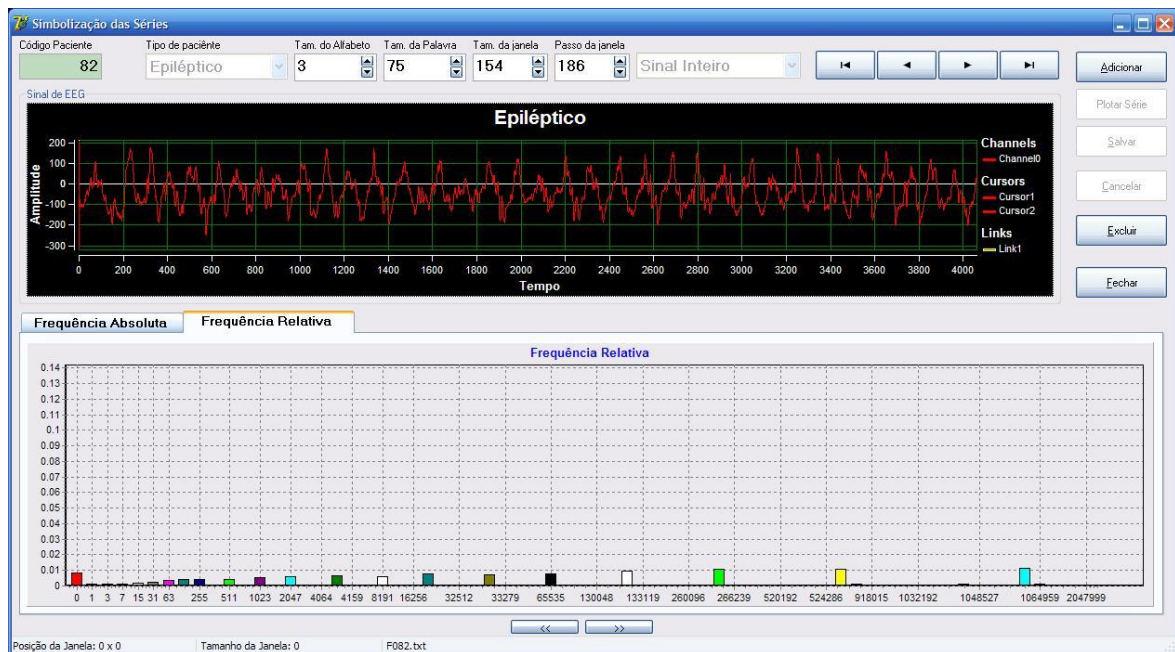


Figura 6-3 – Tela de Simbolização dos Sinais.

Na tela de simbolização de sinais é possível carregar um sinal para o sistema, ajustar parâmetros, efetuar a simbolização e salvar todos os dados provenientes deste sinal.

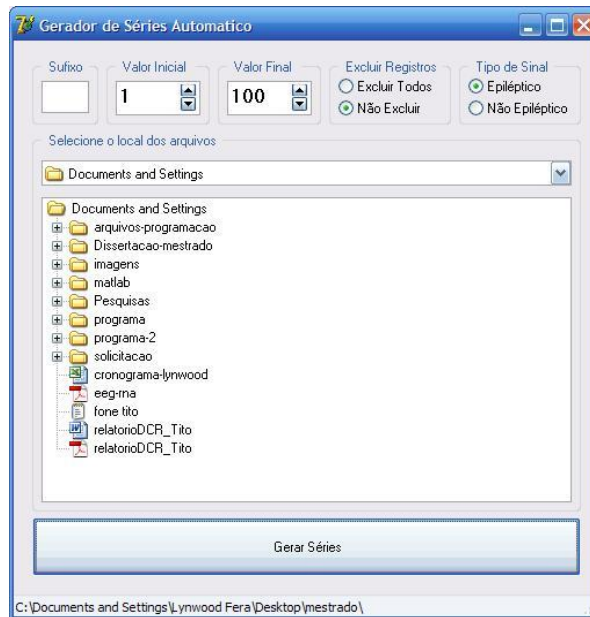


Figura 6-4 – Tela do Simbolizador Automático.

Nesta parte do sistema é possível simbolizar um pacote de sinais em cadeia, caso seja necessário ele irá simbolizar todos salvando na base de dados.



Figura 6-5 – Tela do Otimizador Algoritmo Genético.

Nesta parte do sistema é possível otimizar os valores dos parâmetros para a simbolização, onde através do Algoritmo Genético desenvolvido, tentará encontrar os melhores parâmetros para a simbolização.

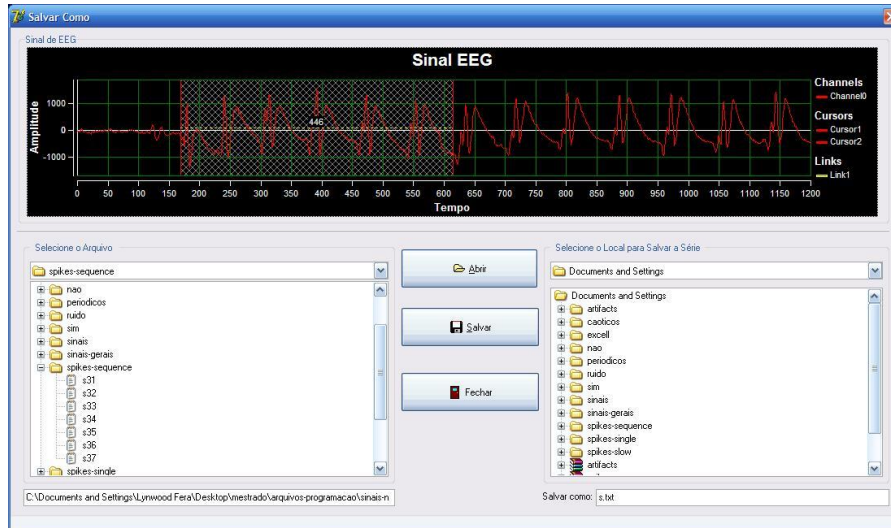


Figura 6-6 – Tela para separar os eventos e salvá-los.

Nesta parte do sistema é possível carregar um sinal, possibilitando o usuário separar as partes que sejam importantes para implementar e ampliar a base de dados do sistema, tanto para otimização quanto para treinamento da rede.

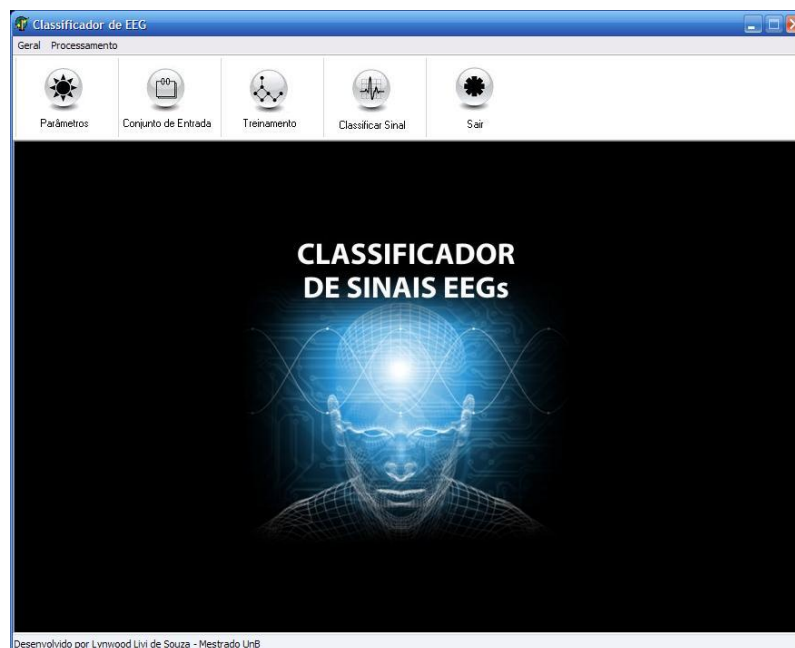


Figura 6-7 – Tela principal do Classificar RNA.

Esta parte do sistema possibilita o acesso às demais partes do sistema Classificador RNA, onde poderá acessar a configuração dos parâmetros da rede, visualizar e ajustar os dados que farão parte do treinamento da rede, acessar o módulo de construção e treinamento da rede e também o módulo de acesso para classificar o sinal.

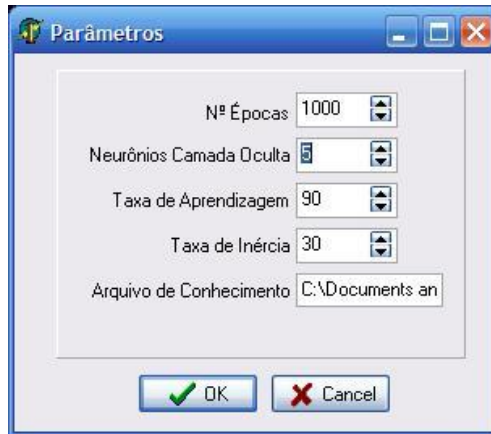


Figura 6-8 – Tela de ajuste dos parâmetros da Rede neural Artificial.

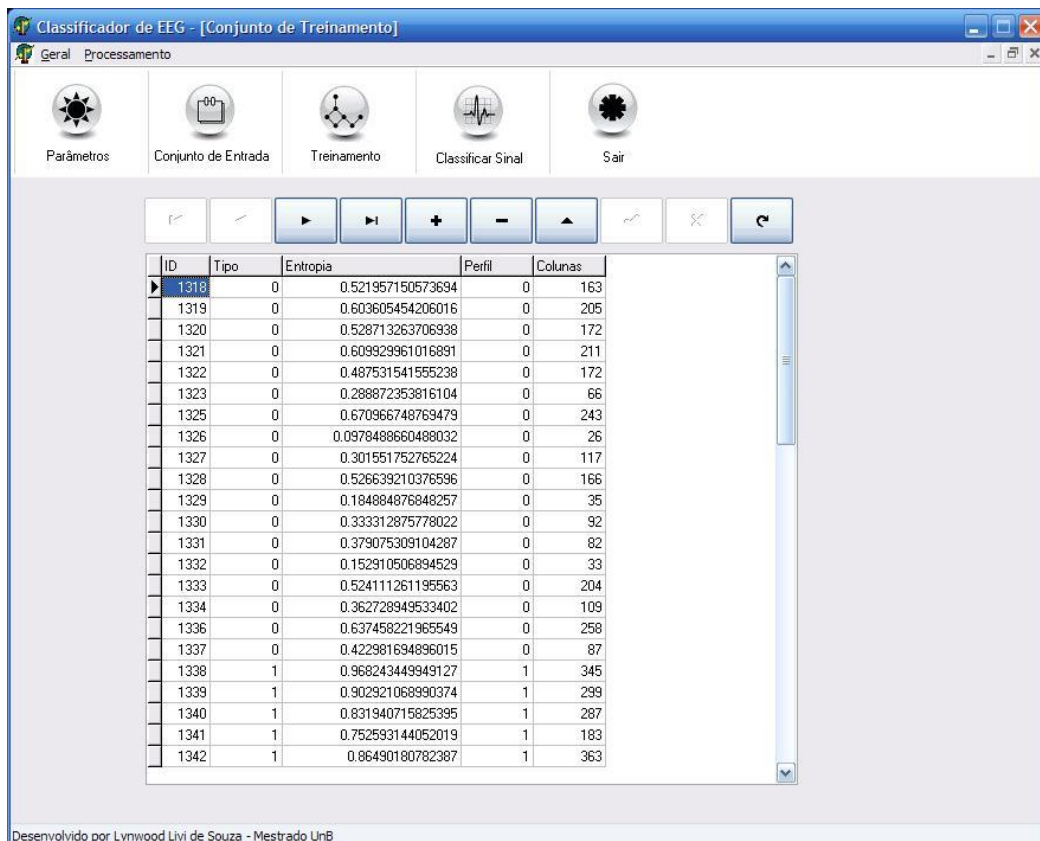


Figura 6-9 – Tela Conjunto de Treinamento.

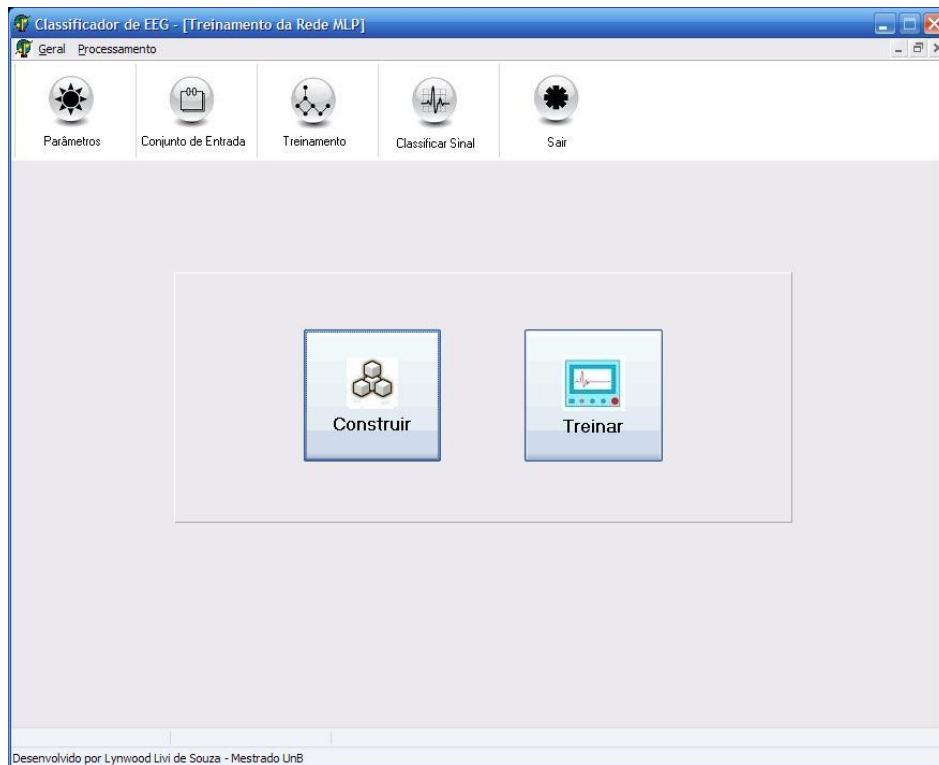


Figura 6-10 – Tela de Construção e Treinamento da Rede Neural Artificial.

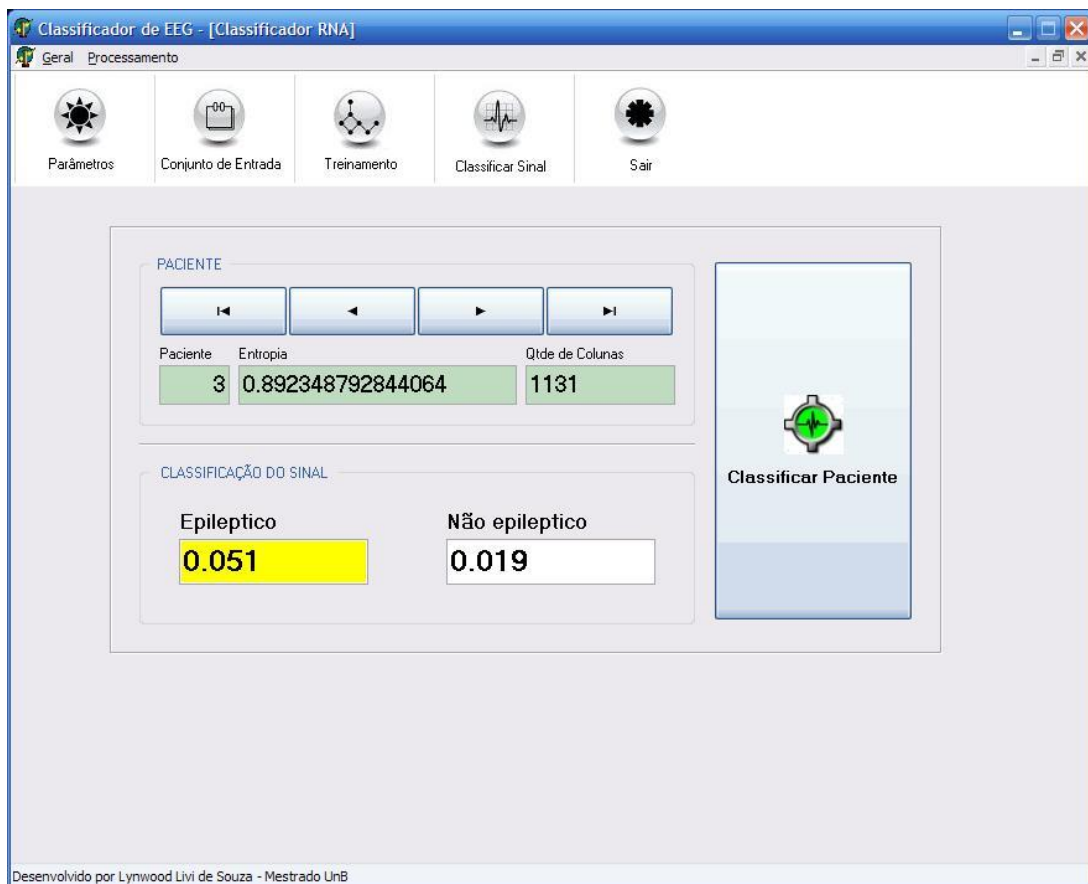


Figura 6-11 – Tela do Classificador.

Esta parte do sistema é capaz de receber o sinal simbolizado e classificar o seu perfil, podendo ser “*Epiléptico*” ou “*Não Epiléptico*”.

6.2. SIMBOLIZAÇÃO DE SINAIS PERIÓDICOS E CAÓTICOS

Por estar apresentando uma nova técnica para detecção automática de eventos epileptiformes em sinais de EEG, também é mostrado como é a dinâmica dos resultados utilizando sinais periódicos e sinais caóticos, bastante usados na literatura. Esta comparação auxilia na distinção e na comparação dos resultados entre os sinais utilizados a seguir, demonstrando a eficácia deste novo método para realizar a detecção dos eventos epiléticos em exames de EEG.

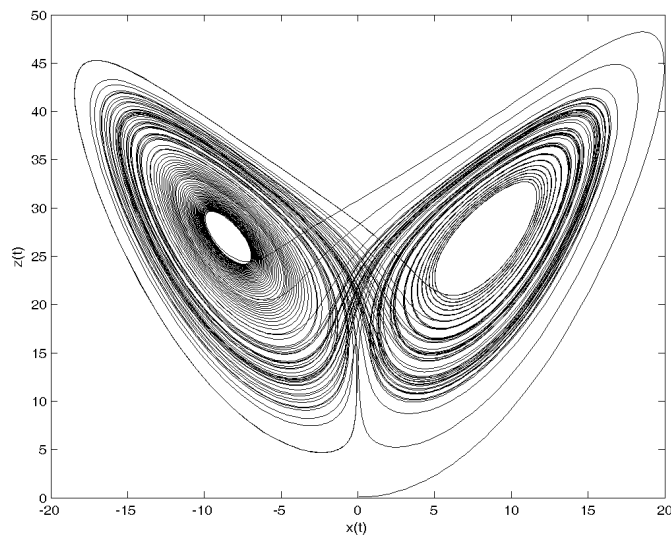


Figura 6-12 – Figura de um sinal caótico.



Figura 6-13 – Figura de um sinal periódico.

Para realizar estes testes foram utilizadas vinte amostras de sinais periódicos e vinte amostras de sinais caóticos. Através do sistema desenvolvido, aplicamos para estes sinais a realização da simbolização por séries temporais. Neste contexto, torna-se necessária a

obtenção dos melhores parâmetros para a simbolização, onde se possa encontrar os melhores valores que tornem a simbolização mais eficiente para a comparação que diferencie os sinais caóticos dos sinais periódicos.

Para encontrar os melhores valores que diferenciam estes sinais, utilizou-se os Algoritmos Genéticos, por se tratar de um método muito eficiente para otimização. A utilização dos Algoritmos Genéticos visa encontrar os melhores valores para os parâmetros “Tamanho da Palavra”, “Tamanho do Alfabeto”, “Tamanho da janela” e “Tamanho do passo da janela”, que são os principais parâmetros utilizados para a simbolização das séries temporais.

Para otimização utilizando os Algoritmos Genéticos, foram definidos os seguintes valores descritos na Tabela 6.1 para cada operador genético.

OPERADORES GENÉTICOS	VALORES
TAMANHO DO CROMOSSOMO	2
TAMANHO DA POPULAÇÃO	100
TAXA DE CROSSOVER	0.80
TAXA DE MUTAÇÃO	0.90
PONTOS DE CROSSOVER	2
INDÍVIDUOS NO ELITISMO	5
NÚMERO DE GERAÇÕES	20
TIPO DE SELEÇÃO	RANK
TIPO DE OTIMIZAÇÃO	MAXIMIZAÇÃO

Tabela 6.1 – Valores dos operadores genéticos.

Na otimização é inserida toda a simbolização das séries, para que assim possa fazer todas as comparações entre as diversas gerações, mantendo sempre os melhores indivíduos para as gerações futuras.

Dentro da simbolização são gerados histogramas relativos às frequências relativas e absolutas referentes aos sinais analisados, desta forma é possível observar a dinâmica para cada sinal.

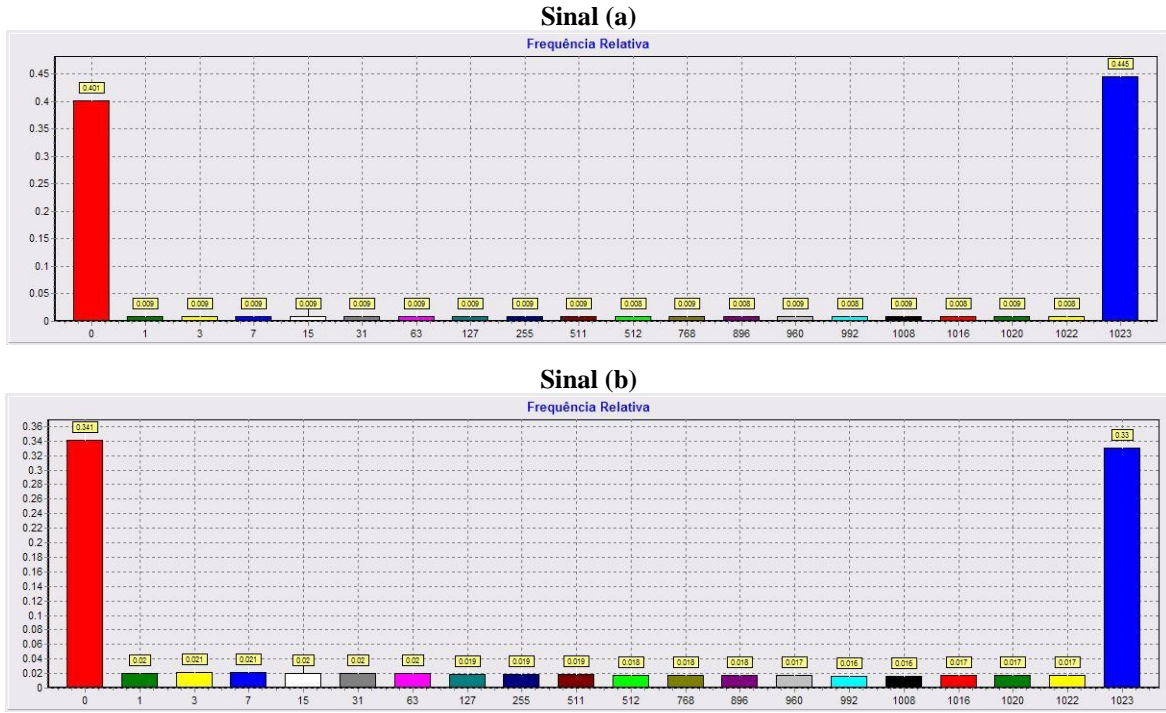
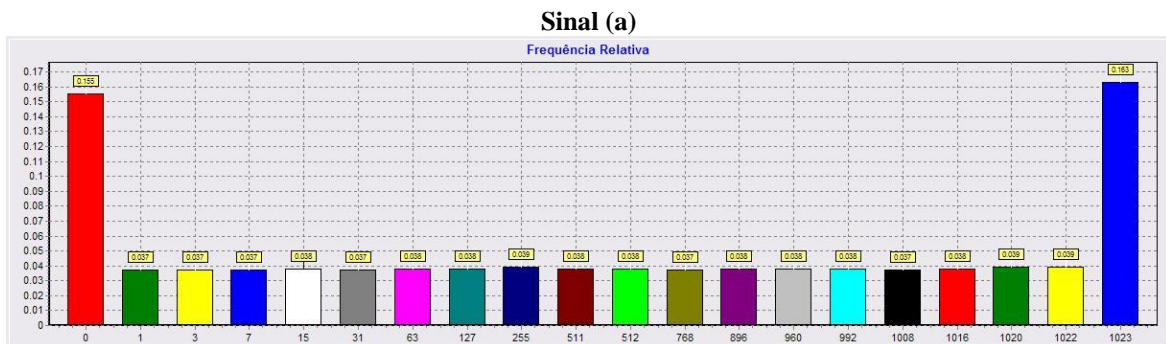


Figura 6-14 – Frequência relativa de sinal caótico utilizando parâmetros de simbolização aleatórios.

Na Figura 6.14 estão demonstradas as frequências relativas de dois sinais caóticos e na Figura 6.15 as frequências relativas de dois sinais periódicos, onde foram utilizados valores aleatórios para os parâmetros da simbolização das séries temporais



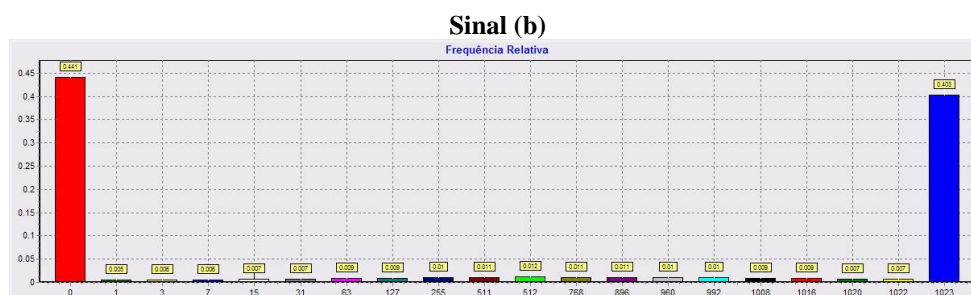


Figura 6-15 – Frequência relativa de dois sinais periódicos utilizando parâmetros de simbolização aleatórios.

A partir dos valores dos histogramas acima, onde foram utilizados parâmetros aleatórios, observa-se que alguns valores não permitem a distinção dos sinais por meio da simbolização. Diante de tais características observou-se a necessidade de implementar uma função estatística para quantificar a diferenciação entre os histogramas. Para análise estatística dos histogramas, utilizamos a Entropia de Shannon, aplicada dentro do Algoritmo Genético.

Utilizando a técnica descrita foi possível encontrar com a otimização os parâmetros que proporcionaram a maior diferença entre os tipos de sinais em tela. Dentre os parâmetros que eram buscados, apresenta-se os valores ao final das gerações:

PARÂMETROS	VALORES
TAMANHO DA PALAVRA	93
TAMANHO DO ALFABETO	3
TAMANHO DA JANELA	94
PASSO DA JANELA	49

Tabela 6.2 – Valores otimizados pelo Algoritmo Genético.

Utilizando os parâmetros especificados na Tabela 6.2, aplicados aos vinte sinais caóticos e os vinte sinais periódicos, realizou-se a soma de todos os sinais caóticos e a soma de todos os sinais periódicos e ao final aplicou-se a subtração destas somas onde se encontrou a melhor aptidão entre várias gerações de indivíduos, obtendo-se como a maior aptidão o valor de 19,68.

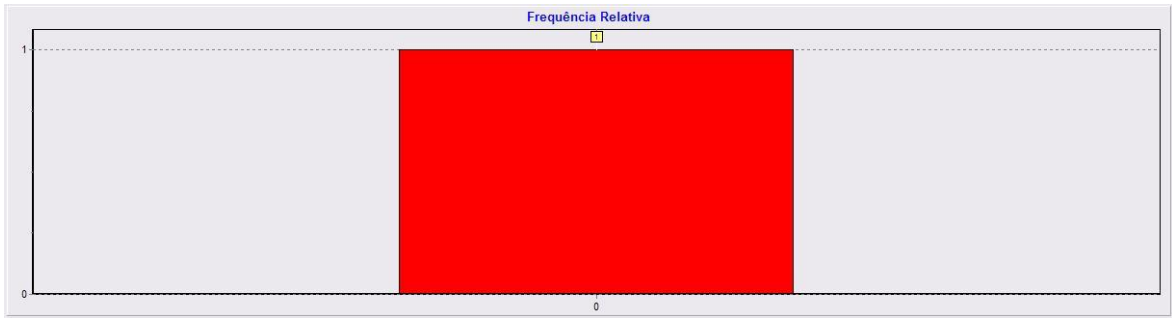


Figura 6-16 – Frequência relativa do sinal Caótico após a simbolização com os valores dos parâmetros otimizados pelo Algoritmo Genético.

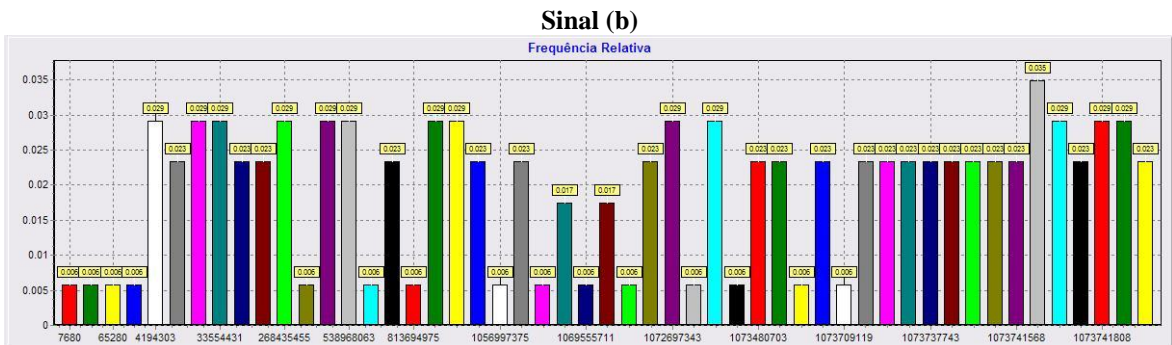
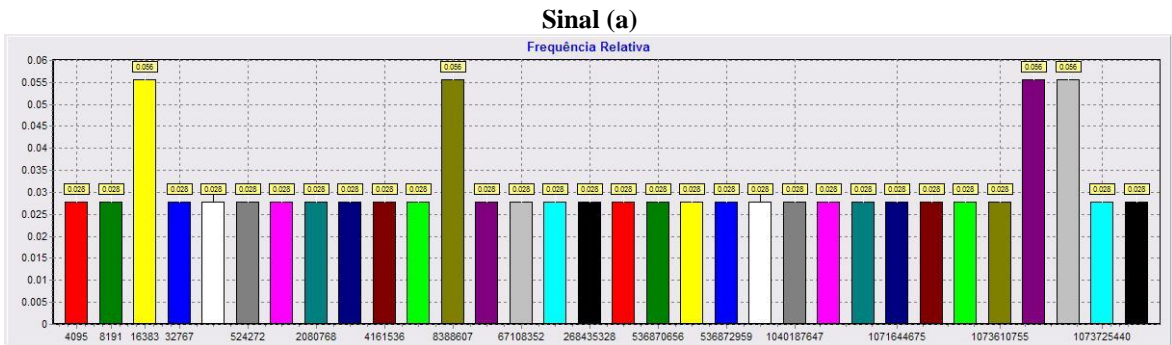


Figura 6-17 – Frequência relativa dos Sinais Periódicos após a simbolização com os valores dos parâmetros otimizados pelo Algoritmo Genético.

Portanto, podemos observar uma grande eficácia utilizando os métodos citados neste trabalho, onde conseguimos diferenciar visualmente através dos histogramas sinais caóticos dos periódicos.

6.3. INFLUÊNCIA DO RUÍDO NA SIMBOLIZAÇÃO

Neste tópico é demonstrado a influência do ruído na simbolização dos sinais Caóticos e Periódicos. Desta forma pega-se os sinais já existentes em nossa base de dados e adicionamos ruído aleatório ao sinal, onde enfim o sinal passou a possuir componentes de alta frequência.

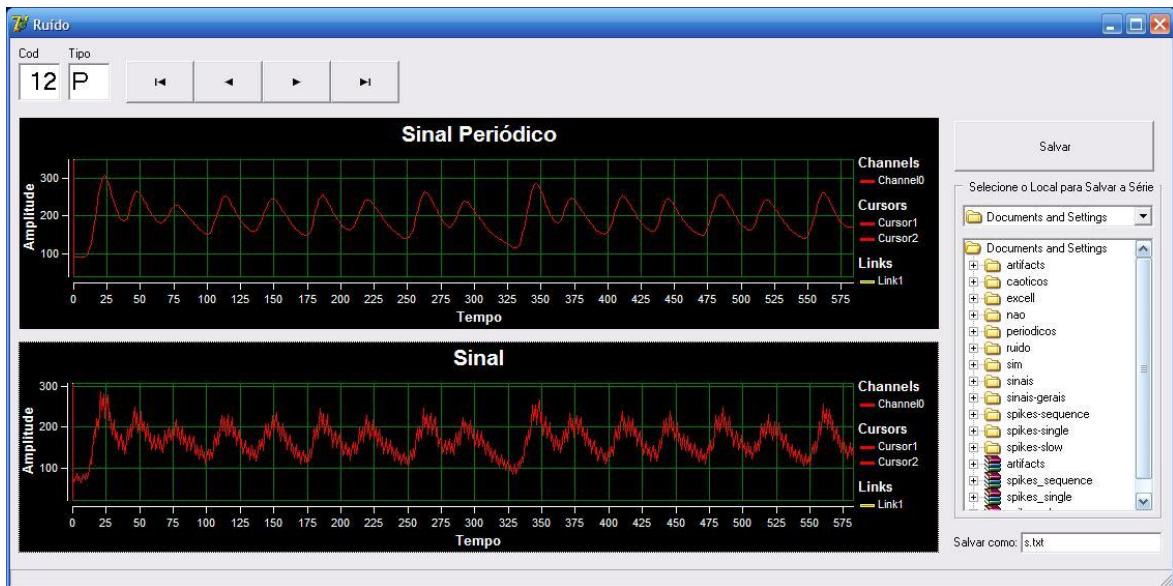


Figura 6-18 – Tela do programa demonstrando o sinal Periódico normal e abaixo o sinal Periódico adicionado de ruído aleatório.

Após salvar os sinais com a presença de ruídos, em seguida foi realizada a simbolização para comparação dos sinais Caóticos com ruído, e sinais Periódicos com ruído.

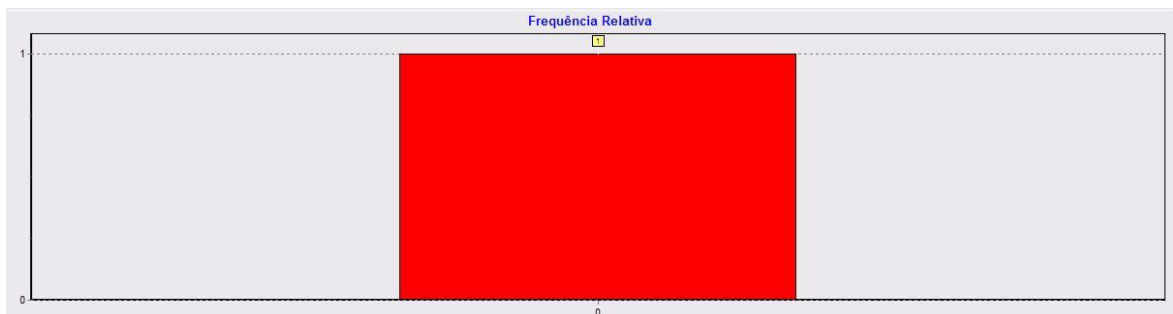


Figura 6-19 – Frequência relativa de um sinal Periódico com ruído após a simbolização.

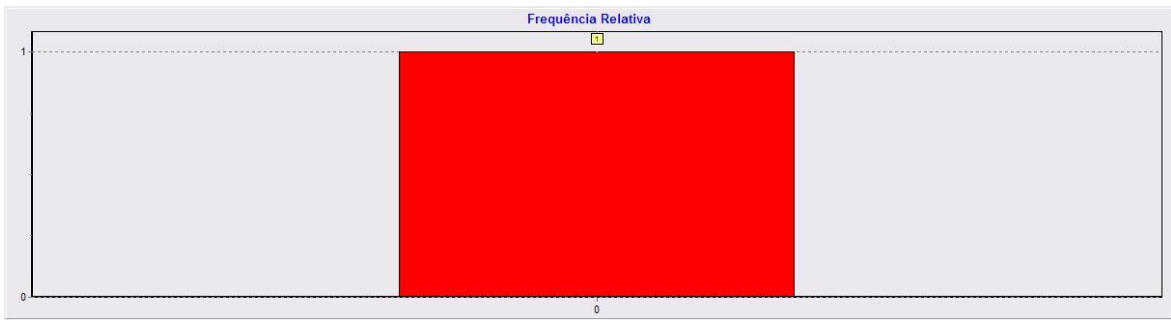


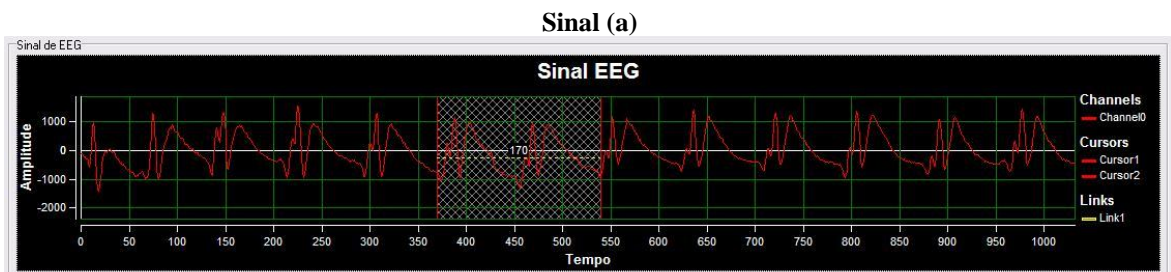
Figura 6-20 – Frequência relativa de um sinal Caótico com ruído após a simbolização.

Observando a Figura 6.19 e Figura 6.20, fica evidente a grande influência que um ruído ocasiona em um sinal caótico e periódico para a simbolização das séries, onde os sinais de forma normal eram visíveis suas diferenças e ao adicionar ruído se torna impossível uma diferenciação entre eles utilizando os parâmetros com os valores otimizados pelo Algoritmo Genético.

6.4. SIMBOLIZAÇÃO DE EEG NORMAIS E EPILÉPTICOS

Como está proposto neste projeto, iremos simbolizar os sinais EEG com resultados normais e EEG com eventos epiléticos, desta forma poderemos demonstrar através da dinâmica simbólica uma forma de detectar automaticamente com um grau de certeza muito grande a diferenciação entre os sinais.

A princípio, o primeiro passo adotado, foi a separação de eventos epiléticos para a simbolização, os quais são eventos característicos que o sistema utilizará posteriormente para comparação, formando a base de dados com duas características distintas.



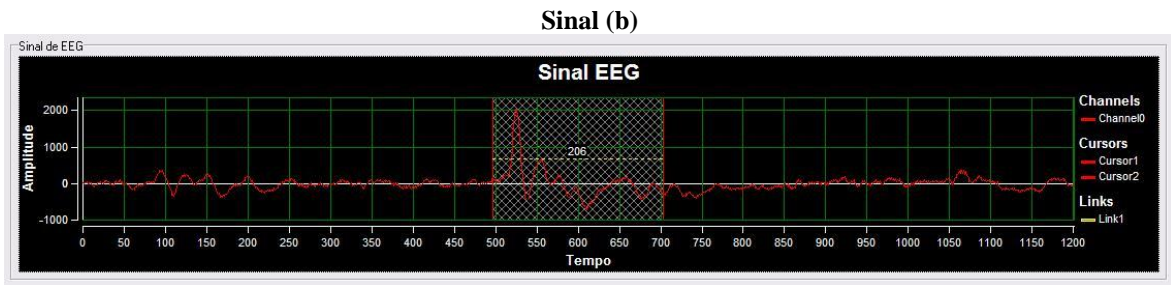


Figura 6-21 – Separação dos eventos para formação da Base de Dados.

Realizado esse processo onde separamos os eventos com características específicas de epilepsia, salvamos estes sinais para aplicarmos a simbolização, onde através da simbolização das séries temporais vamos abstrair características claras que diferenciem ambos os sinais.

Para uma simbolização primária, utilizou-se parâmetros arbitrários para que fosse possível observar os resultados, e comparar se há uma diferenciação entre os sinais. A Tabela 6.3 lista os parâmetros utilizados de forma arbitrária para esta primeira análise:

PARÂMETROS	VALORES
TAMANHO DA PALAVRA	5
TAMANHO DO ALFABETO	3
TAMANHO DA JANELA	15
PASSO DA JANELA	5

Tabela 6.3 – Valores dos parâmetros arbitrários na simbolização das séries temporais.

Utilizando os valores descritos na Tabela 6.3, podemos observar a dinâmica de seus resultados através das frequências relativas de cada valor gerado pela simbolização:

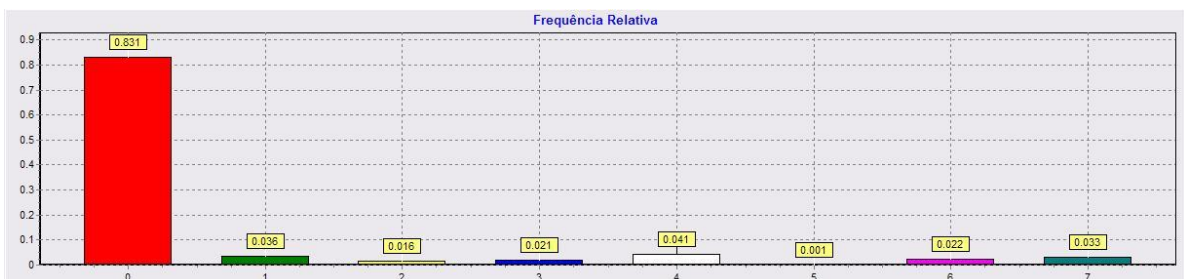


Figura 6-22 – Frequência Relativa dos sinais sem a presença de eventos epiléticos utilizando os valores da Tabela 6.3.

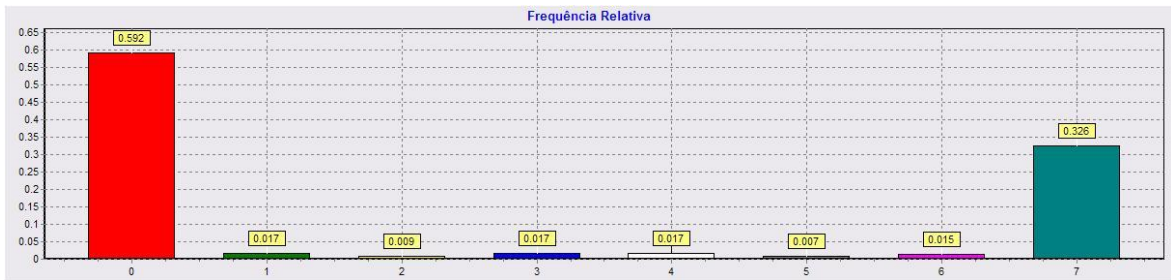


Figura 6-23 – Frequência Relativa dos sinais com a presença de eventos epilépticos utilizando os valores da Tabela 6.3.

Pode-se observar que tanto a frequência relativa dos sinais com eventos epilépticos quanto a frequência relativa dos sinais não epilépticos utilizando os parâmetros de forma arbitrária não se é possível diferenciar a dinâmica de cada valor das frequências de acordo com cada sinal. Portanto, para encontrar parâmetros que realmente realize a simbolização das series e encontre uma dinâmica que diferencie tais sinais, foram utilizados o Algoritmo Genético para realizar a otimização dos parâmetros necessários, de modo a permitir a simbolização das séries. Deste modo, é possível encontrar valores para os parâmetros que melhor diferenciam o resultado das frequências entre os sinais com eventos epilépticos e não epilépticos.

Para otimização utilizando os Algoritmos Genéticos, foram definidos os seguintes valores descritos na Tabela 5.4 para cada operador genético.

OPERADORES GENÉTICOS	VALORES
TAMANHO DO CROMOSSOMO	4
TAMANHO DA POPULAÇÃO	100
TAXA DE CROSSOVER	0.80
TAXA DE MUTAÇÃO	0.90
PONTOS DE CROSSOVER	2
INDÍVIDUOS NO ELITISMO	5
NÚMERO DE GERAÇÕES	30
TIPO DE SELEÇÃO	RANK
TIPO DE OTIMIZAÇÃO	MAXIMIZAÇÃO

Tabela 6.4 – Valores dos operadores genéticos utilizados para otimização.

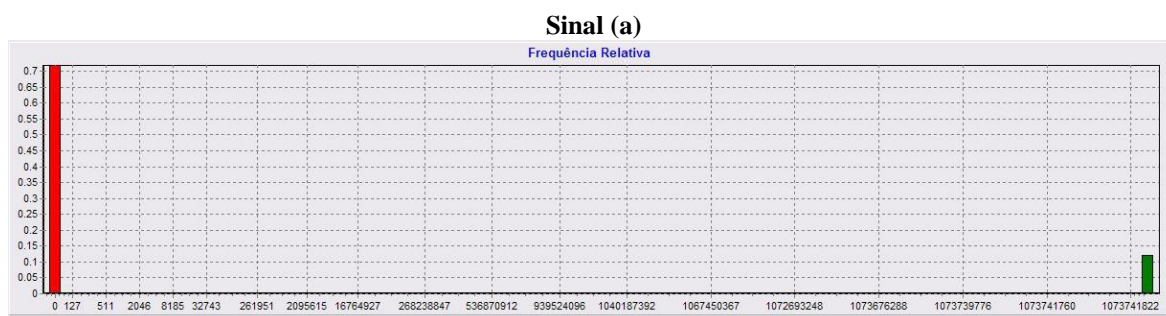
Utilizando estes valores para os operadores genéticos, buscamos otimizar os parâmetros para a simbolização das séries, desta forma aplicamos juntamente a entropia de shannon, onde, somamos a aptidão de cada indivíduo com eventos epiléticos e subtraímos com a soma das aptidões dos indivíduos com sinais normais.

A partir da otimização, juntamente com uma ferramenta estatística para encontrar a aptidão de cada indivíduo, somando todas as aptidões separadamente de cada tipo de sinal e subtraindo-os, encontrou-se parâmetros que melhor diferenciaram os sinais, tendo o melhor indivíduo com a aptidão 9,20095650462122. Os parâmetros que foram utilizados para que se encontrasse esse indivíduo estão descritos na Tabela 6.5:

PARÂMETROS	VALORES
TAMANHO DA PALAVRA	75
TAMANHO DO ALFABETO	3
TAMANHO DA JANELA	154
PASSO DA JANELA	183

Tabela 6.5 – Valores dos parâmetros otimizados na simbolização das séries temporais utilizando o Algoritmo genético.

Através dos parâmetros otimizados, pode-se aplicar a simbolização das series temporais, como também perceber as freqüências de cada sinal, tanto os eventos epiléticos como não epiléticos, conforme freqüências abaixo:



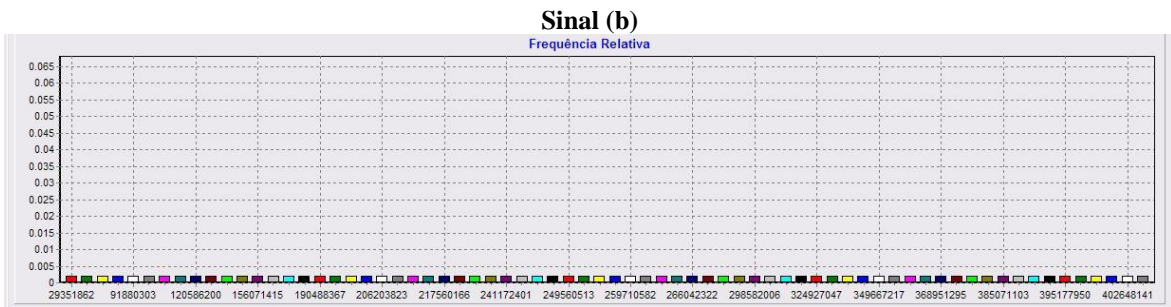


Figura 6-24 – Frequência relativa dos Sinais Normais após a simbolização utilizando os parâmetros da Tabela 6.5.

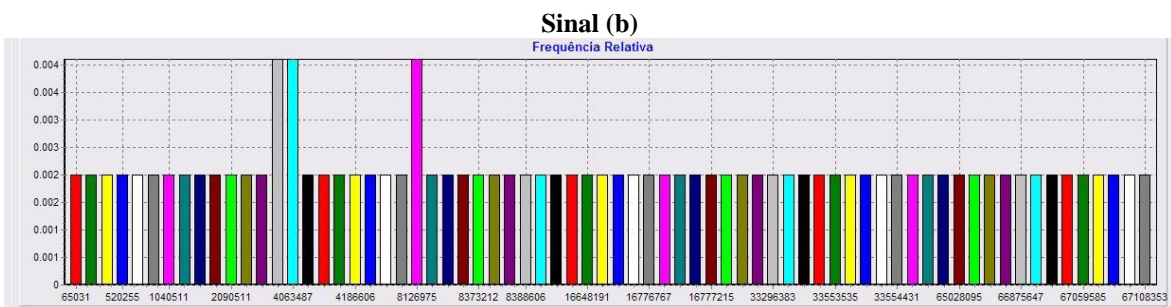
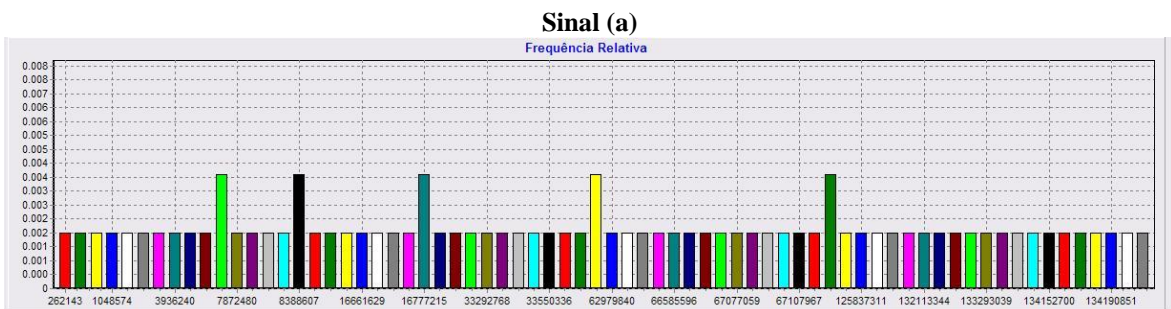


Figura 6-25 – Frequência relativa dos Sinais Epilépticos após a simbolização utilizando os parâmetros da Tabela 6.5.

Portanto, fica evidente a eficácia da utilização das técnicas de otimização e estatística, descritas acima para a simbolização das séries temporais. Em outras palavras, utilizando os valores encontrados pelo Algoritmo Genético, foi possível encontrar parâmetros que ao final da simbolização, apresentou valores das colunas dos histogramas, capazes de diferenciar os dois tipos de sinais analisados.

6.5. TREINAMENTO DA REDE NEURAL PARA A CLASSIFICAÇÃO

Após o término da simbolização dos sinais, será construído e treinado a rede neural, na qual foi descrita na metodologia, para que possa classificar de forma automática os sinais.

Após definição dos perfis que devem ser classificados pela rede, deve-se especificar as entradas da rede MLP. A sugestão, de acordo com a descrição é alimentar a rede com o valor da *Entropia de Shannon* aplicada a cada sinal juntamente com a quantidade de colunas apresentada por cada frequência do sinal. Assim definimos nossa rede com dois neurônios de entrada, uma camada oculta contendo três neurônios, dois neurônios de saída, cinco mil o número de épocas para iteração de treinamento da rede MLP, oito por cento o número da taxa de aprendizagem e taxa de inércia de três por cento.

CONSTRUÇÃO DA REDE MLP	
NÚMERO DE ÉPOCAS	5000
NEURÔNIOS DA CAMADA OCULTA	3
TAXA DE APRENDIZAGEM	0,08
TAXA DE INÉRCIA	0,03
ARQUIVO DE CONHECIMENTO	Perfil.MLP

Tabela 6.6 – Parâmetros para construção da rede MLP.

Definindo a arquitetura e o valor dos parâmetros de nossa rede MLP, vamos especificar o conjunto de treinamento a ser utilizado pela rede.

CONJUNTO DE TREINAMENTO			
COD	ENTROPIA	COLUNAS	PERFIL
1	0.521957150573694	163	0
2	0.603605454206016	205	0
3	0.528713263706938	172	0
4	0.609929961016891	211	0
5	0.487531541555238	172	0
6	0.288872353816104	66	0
7	0.670966748769479	243	0

8	0.0978488660488032	26	0
9	0.301551752765224	117	0
10	0.526639210376596	166	0
11	0.184884876848257	35	0
12	0.333312875778022	92	0
13	0.379075309104287	82	0
14	0.152910506894529	33	0
15	0.524111261195563	204	0
16	0.362728949533402	109	0
17	0.637458221965549	258	0
18	0.422981694896015	87	0
19	0.603605454206017	205	0
20	0.637458221965549	258	0
21	0.670966748769479	243	0
22	0.968243449949127	345	1
23	0.902921068990374	299	1
24	0.831940715825395	287	1
25	0.752593144052019	183	1
26	0.86490180782387	363	1
27	0.863232739586893	323	1
28	0.918784248344332	331	1
29	0.998396382409292	477	1
30	0.996153895935167	463	1
31	0.915984338266393	389	1
32	0.936317159149031	420	1
33	0.993005103174491	458	1
34	0.990915090191592	460	1
35	0.965933537866117	446	1
36	0.879430540164009	385	1
37	0.999371075837742	485	1
38	0.999746643106741	488	1
39	0.995222988675676	463	1
40	0.994428580001243	473	1
41	0.983383402020252	445	1
42	0.758681775260041	814	1
43	0.881176765211752	1044	1
44	0.613245781194914	521	1

Tabela 6.7 – Conjunto de treinamento para rede MLP.

Na Tabela 6.7, observamos a coluna *Entropia* com os valores da *Entropia de Shannon* dos valores para cada sinal observado, na coluna *Colunas* temos a quantidade de colunas existente em cada frequência e na coluna *Perfil* definimos o perfil de cada sinal o qual representa. Dentro do perfil especificamos “0” para sinais sem a presença de eventos epiléticos e “1” para sinais com presença de eventos epiléticos, onde alimentamos a

rede com vinte e um valores constatados não epiléticos e vinte e três com valores constatados epiléticos.

A partir dos modelos citados acima, realizamos a parte de construção e treinamento da rede, ao final do treinamento o erro total foi de 0.072321. Em seguida é realizada a sessão de teste, a rede foi carregada com o arquivo de conhecimento salvo para efetuar as classificações conforme padrões apresentados à sua entrada. O intervalo definido para os neurônios de saída permitem que trabalhem com um conceito aproximado de probabilidade.

Após a execução dos testes da rede, encontraram-se os seguintes resultados:

REDE NEURAL MLP	
DESCRIÇÃO	VALORES
ACERTOS	96%
ERROS	4%
VERDADEIROS POSITIVOS	96,8%
FALSOS NEGATIVOS	3,2%
VERDADEIROS NEGATIVOS	88%
FALSOS POSITIVOS	12%
ESPECIFICIDADE	88%
SENSIBILIDADE	96,8%
SELETIVIDADE POSITIVA	98,7%
SELETIVIDADE NEGATIVA	73,3%
PRECISÃO	92,4%

Tabela 6.8 – Resultados dos teste obtidos neste trabalho.

A base de dados utilizada no treinamento e nos testes deste projeto são oriundas do Departamento de Epileptologia da Universidade de Bonn – Alemanha (Universidade de Bonn) e também do grupo de estudos e desenvolvimento do *software Wavelet Analysis of Epileptic Spikes* (EEG SPIKES DETECTION). A Tabela 6.9 aborda sobre os dados utilizados:

DESCRIÇÃO	VALOR
NÚMERO DE SINAIS	550
TAXA DE AMOSTRAGEM	173,61 Hz e 240 Hz
LARGURA DE BANDA ESPECTRAL	0,5 Hz a 85 Hz
FILTRO PASSA-BAIXA	40 Hz
TAMANHO DA JANELA	4098 pontos

Tabela 6.9 – Informações dos arquivos utilizados para compor da base de dados

7. CONCLUSÕES E DISCUSSÕES

A principal motivação para realizar este trabalho foi colaborar com o desenvolvimento das pesquisas na área da tecnologia, à medida que foi possível desenvolver e demonstrar uma nova técnica de identificação e caracterização de eventos epileptiformes em sinais de EEG, utilizando a dinâmica simbólica. Ou seja, através dos sinais dos exames de EEG observou-se toda a dinâmica existente entre os sinais, com presença de eventos de epilepsia, aplicando em seguida as redes neurais artificiais para auxiliar na classificação. Depois de observado a dinâmica destes sinais, aparecem dois resultados: sinal epilético ou sinal não epilético. Dentro das pesquisas realizadas, observou-se na literatura alguns resultados de trabalhos que utilizaram diversas técnicas para o mesmo propósito.

Nos estudos de Zwe-Lee (2004) aplicou-se uma metodologia baseada na transformada de Wavelet discreta aliada com um classificador de redes neurais do modelo probabilístico, onde obteve em seus resultados um valor de 90% de acertos.

Pereira (2003) utilizou o modelo de redes neurais com entradas parametrizadas, onde desta forma obtêm-se os parâmetros observando diretamente o sinal, aplicando-se à rede neural quarenta e cinco parâmetros extraídos do sinal para o reconhecimento da rede que utiliza o modelo Perceptron Multicamadas treinada pelo algoritmo Backpropagation. Esta técnica esbarra em algumas dificuldades como por exemplo, uma profunda análise, capaz de eliminar equívocos relacionados à amplitude de eventos de epilepsia com ruídos e piscadas palpebrais. Contudo, esse método atingiu 85,2% de acertos, 89% de sensibilidade, 90% de especificidade e 89,5% de precisão.

No trabalho de Lui *at al.*, (2002) propôs-se a utilização de vários métodos para o processamento de sinais com diversos estágios como: filtros adaptativos integrados onde em primeira instância tinha o objetivo de separar a parte estacionária da parte não estacionária no sinal. Em seguida a transformada de wavelet para a extração de características; redes neurais artificiais para o reconhecimento de características onde utilizou a topologia de perceptrons com várias camadas; e por fim os sistemas especialistas para separar os artefatos que possam surgir no sinal utilizando regras heurísticas para

decidir se o sinal possui eventos epileptogênicos ou não epileptogênicos. Com essas técnicas foi possível alcançar um total de acertos de 90,0%.

Em Argoud *et. al.*, (2004) foi proposto utilização da Transformada Wavelet para extrair características de eventos com a presença epileptogênica no sinal de exame de EEG, e um grupo de redes neurais artificiais para classificação e distinção dos eventos epileptogênicos dos eventos normais de fundo. O sistema proposto atingiu um desempenho de 80% de acertos.

No mesmo sentido, Patnaik *et al.*, (2008) utiliza wavelet para a extração de características e obtenção dos parâmetros estatísticos a partir dos coeficientes wavelet. Uma rede neural artificial com topologia *feedforward* aliada com o algoritmo de treinamento *backpropagation* para a classificação destas características. Utilizou-se também o algoritmo genético para a escolha do conjunto de treinamento e também para implementar um estágio de pós-classificação utilizando pesos harmônicos para aumentar a precisão. Enfim este trabalho alcançou uma especificidade média de 99,19%, sensibilidade de 91,29% e seletividade de 91,14%.

Observando algumas das principais literaturas sobre detecção e caracterização de eventos epileptiformes em sinais de EEG, estão em média divididas em duas partes: extração de características e classificação destas características. A extração de características o qual é um fator primordial para que haja uma ótima classificação, são utilizadas diversas técnicas, afim de que se encontre a que melhor consiga resolver a enorme dinâmica existente dentro de um sinal de EEG.

Para a classificação destas características a literatura é unânime em salientar a utilização das Redes Neurais Artificiais, devido a sua capacidade de classificar e aprender ao longo de sua aplicação.

Por se tratar de um desafio, tendo em vista que a proposta é resolver o grande problema da extração de características, sugere-se com esta pesquisa, um novo método capaz de dar resolatividade nesta área da ciência.

Neste trabalho foi aplicada a Dinâmica Simbólica para extração de características que evidenciem a presença de eventos epileptiformes, e de modo satisfatório foi possível alcançar os resultados com essa nova técnica. Para tanto, foi utilizada ferramentas de otimização como os Algoritmos Genéticos, a Dinâmica Simbólica para extração de características e as Redes Neurais Artificiais para a classificação destes eventos, chegando a resultados amplamente favoráveis, tendo em vista que alcançamos 96,8% de sensibilidade, 88% de especificidade, 92,4% de precisão e 96% de acertos. São números positivos para pesquisas que envolvem novos métodos de aplicação e que podem atingir parâmetros mais significativos, a partir da constatação deste trabalho.

8. SUGESTÕES PARA TRABALHOS FUTUROS

Levando-se em conta os resultados satisfatórios encontrados neste trabalho, com o método de extração de características utilizando a Dinâmica Simbólica juntamente com um classificador em Redes Neurais Artificiais, convém algumas sugestões para trabalhos futuros, afim que se possa comparar a capacidade de precisão da técnica utilizada:

- O aumento do banco de dados de EEG utilizados: onde a idéia é ter uma capacidade maior de sinais, onde assim o sistema seja capaz de analisar diversos tipos de sinais com inúmeras dinâmicas aumentando sua capacidade de extração e classificação de características.
- Utilizar Algoritmos Genéticos para otimização dos parâmetros da Rede Neural Artificial, a fim de encontrar valores que estejam próximos de uma solução ótima, para que melhor se torne o treinamento da rede.
- Realizar pesquisas para encontrar os principais fatores que influenciam na detecção dos resultados Falsos Positivos (FP), que são resultados que diminuem sensivelmente a capacidade e precisão de um sistema de detecção de eventos epileptiformes.
- Implantar métodos paramétricos e estatísticos a fim de aprimorar a capacidade do sistema de observar maiores números de características apresentadas pelos sinais de EEG.
- Utilizar a Dinâmica Simbólica em outros tipos de sinais como: Eletrocardiograma, exames do aparelho respiratório e exame da avaliação dinâmica da musculatura. Onde podemos contribuir com esta técnica para auxiliar nos resultados de patologias e características observadas através destes exames.

9. REFERÊNCIAS BIBLIOGRÁFICAS

ANDERSON JA & SUTTON JP. A Network of Networks: Computation and Neuro-biology. World Congress on Neural Networks, 1995, 1, 561-568.

ARGOUD, F. I. M.; AZEVEDO, F. M.; NETO, J. M. Sistema de detecção automática de paroxismos epileptogênicos em sinais de eletroencefalograma. Revista Controle & Automação, Vol.15 no. 4, Outubro, Novembro e Dezembro 2004.

ATLAS, L.; DUHAMEL, P. Recent developments in the core of digital signal processing. Signal Processing Magazine, IEEE, v. 16, n. 1, p. 16-31, 1999.

AZEVEDO, F. M. Algoritmos genéticos em redes neurais artificiais. V Escola de Redes Neurais. ITA. São José dos Campos – SP. 19 de Julho de 1999.

BARRETO, JORGE M. Introdução às Redes Neurais Artificiais. Anais V Escola Regional de Informática da SBC Regional Sul, 5 a 10 de maio de 1997. Páginas 41 - 71.

BORN, R. S. Filtros adaptativos aplicados a sinais biomédicos. Monografia, Universidade Federal de Pelotas, 2000.

BUTTON, V. L. S. N. Eletroencefalógrafos e monitores de EEG. Campinas: Editora Unicamp, 2000.

CABRA, E. F & Grupo Comunicação homem Máquina. Redes neurais Artificiais – um curso técnico e prático para engenheiros e cientistas. Edição dos autores, 1999, 1ª ed.

CAMINAL, P., GIRALDO, B., ZABALETA, H., VALLVERDÚ, M., BENITO, S., BALLESTEROS, D., LOPEZ-RODRIGUEZ, L., ESTEBAN, A., BAUMERT, M. e VOSS, A. "Joint Symbolic Dynamic Analysis of Cardiorespiratory Interactions in Patients on Weaning Trials" In: Proceedings of the 27th International Conference of the IEEE – Engineering in Medicine and Biology Society, pp. 4576-4579, Shanghai, Jan. 2005a.

CAMINAL, P., MATEU, J., VALLVERDU, M., GIRALDO, B., BENITO, S. e VOSS, A. "Estimating Respiratory Pattern Variability By Symbolic Dynamics", *Methods of Information in Medicine*, v. 43, n. 1 (Jan), pp. 22-25. 2004.

CAMINAL, P., VALLVERDU, M., GIRALDO, B., BENITO, S., VAZQUEZ, G. e VOSS, A. "Optimized Symbolic Dynamics Approach for the Analysis of the Respiratory Pattern", *IEEE Transactions on Biomedical Engineering*, v. 52, n. 11 (Nov), pp. 1832-1839. 2005b.

CÂNDIDO, J. D. "Métodos computacionais para identificação de padrões e modelos de previsão". 2005.

CHAGAS, Sandro Luiz das. Classificação de sinais de eletroencefalograma usando máquinas de vetores suporte. Dissertação de Mestrado em Engenharia Elétrica, Universidade Presbiteriana Mackenzie, São Paulo, 2009.

COOLEY, J. W.; TUKEY, J. W. An algorithm for the machine calculation of complex fourier series. *Math of Computation*, v. 19, p. 297-301, 1965.

CROSS SS, HARRISON RF, KENNEDY R. Introduction to Neural Networks. *Lancet*. 346:1075-79. 1995.

CRUZ, C. P. T., Medidas e análise estatísticas em sinais bioelétricos sobre pontos de acupuntura. Dissertação de Mestrado. Universidade Federal do Rio Grande do Norte. Natal, 2007.

DARWIN, C., "A origem das espécies". Rio de Janeiro: Ediouro. 2004.

DAW, C. S., Finney, C.E. e Tracy, E.R., A review of symbolic analysis of experimental data, *Review of Scientifics Instruments*, vol. 74, pp. 916-930, 2003.

DENIS, M. Prática da Medicina do Sono. Rua do Estilo Barroco, 755, São Paulo SP: Editora BYK, 1999.

DIAS, J. S.; BARRETO, J. M., “Algoritmo genético: inspiração biológica na solução de problemas - uma introdução”, Revista Marítima Brasileira - Suplemento Especial, Pesquisa Naval, nº 11, p. 105-128. 1998.

EEG SPIKE DETECTION. Wavelet Analysis of Epileptic Spikes, banco de dados de EEG. Disponível em: <http://eegspike.republika.pl/signals.html>. Acessado em: 31 de janeiro de 2010.

FOGEL, D.B., “Evolutionary Computation”, IEEE Press: Piscataway, NJ. 1995.

FURTADO, J.C.; LORENA, L. A. N., “Algoritmo genético construtivo na otimização de problemas combinatórios de agrupamentos”, III Oficina de cortes e empacotamento, lac.inpe.br. 1998.

GUYTON, A. C. Tratado de fisiologia médica. Rio de Janeiro: Guanabara Koogan, 1997.

HAYKIN, S. Neural Networks: a Comprehensive Foundation. 2nd edition. New Jersey, Prentice-Hall, 1999.

HESE, P. V; BOON, P.; VONCK, K.; LEMAHIEU, I. & WALLE, R. V. D. A new method for detection and source for detection and source analysis of eeg spikes. Engineering in Medicine and Biology Society, vol. 3, no. 14, Sept. 2003.

HOLLAND, J.H. Adaptation in Natural and Artificial Systems, Cambridge, Massachusetts: MIT Press, 1975.

KUGLER, M., Uma contribuição ao desenvolvimento de interfaces cérebro-computador utilizando potenciais visualmente evocados. Dissertação de Mestrado. Centro Federal de Educação Tecnológica do Paraná. Curitiba, 2003.

LIND, D. e MARCUS, B. An Introduction to Symbolic Dynamics and Coding. 2 ed., Cambridge University Press. 1996.

LOPES, C. D. Análise de sinais de eeg utilizando a transformada wavelet discreta e as redes neurais artificiais. Dissertação de Mestrado em Engenharia elétrica, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2005.

LUI, H. S.; ZHANG, T. & YANG, F. S. A Multistage, Multimethod Approach for Automatic Detection and Classification of Epileptiform EEG. IEEE-Transactons on biomedical engineering, vol. 49, no. 12, December 2002.

MAESTRI, R., PINNA, G.D., ACCARDO, A., ALLEGRINI, P., BALOCCHI, R., DÁDDIO, G., FERRARIO, M., MENICUCCI, D., PORTA, A., SASSI, R., SIGNORINI, M.G., LA ROVERE, M.T., e CERUTTI, S. “Nonlinear Índices of Heart Rate Variability in Chronic Heart Failure Patients: Redundancy and Comparative Clinical Value”, Journal of Cardiovascular Electrophysiology, v. 18, n. 4 (Abr), pp. 425-433. 2007.

MELLO, R. G. T., OLIVEIRA, L. F., JANDRE, F. C. e NADAL, J. “Correlação Linear e Dinâmica Simbólica para Detecção da Antecipação no Controle Postural Ortostático”: In: Anais do XX Congresso Brasileiro de Engenharia Biomédica, pp. 282-285, São Pedro - SP, 2006.

MICHALEWICZ, Z. (1996), “Genetic algorithm + data structures = evolution programs”, 3º ed. Springer-Verlag.

PACHECO, R. Sistema para Gerenciamento e Processamento de Sinais Eletroencefalográficos. Trabalho apresentado como requisito parcial à obtenção do título de Bacharel em Informática, UFPEL, Pelotas, 1999.

PARREIRA, F. J. Detecção de crises epilépticas a partir de sinais eletroencefalográficos. Tese de doutorado, Universidade Federal de Uberlândia, Uberlândia, 2006.

PATNAIK, L. M.; MANYAM, O. K. Epileptic EEG detection using neural networks and post-classification. Computer methods and programs in biomedicine. vol. 91, Feb. 2008.

PEREIRA, M. C. V. Tratamento de Sinais Bioelétricos para Processamento por Redes Neurais Artificiais. Tese de Doutorado, Instituto de Engenharia Biomédica, Universidade Federal de Santa Catarina, Florianópolis, 2003.

PEREIRA, W. F. Perspectivas Clínico-Terapêuticas baseadas na Engenharia Neural. Tese de Doutorado. Universidade Federal de Uberlândia, Uberlândia, 2004.

PORTILHO, M. C. C. C., CASTRO, D. M. Uma visão geral sobre redes neurais artificiais. Trabalho de Graduação. Universidade Estadual do Oeste do Paraná. Foz do Iguaçu, 2006.

PSIQUIATRIA GERAL. O que devo saber sobre epilepsia. Disponível em: http://www.psiquiatriageral.com.br/epilepsia/manual_epilepsia.htm. Acessado em: 12 de janeiro de 2010.

REGGIA JA, Neural Computation in Medicine. *Artif Intell Med*, 1993, 5, 143-157.

RODRIGUES, M. B. Desenvolvimento de um Instrumento Virtual para Aquisição de Análise de Sinais Bioelétricos. Dissertação de Mestrado em Engenharia Elétrica, Universidade Federal de Santa Catarina, 1997.

SOARES, Pedro Paulo da Silva; NADAL, Jurandir. Aplicação de uma Rede Neural Feedforward com Algoritmo de Levenberg-Marquardt para Classificação de Alterações do Segmento ST do Eletrocardiograma. IV Congresso Brasileiro de Redes Neurais, Julho 20-22, 1999 - ITA, São José dos Campos – SP.

SRINIVAS, M.; PATNAIK, L.M. (1994), “Genetic algorithms: A survey”, IEEE.

STEARNS, S. C. “Evolução: uma introdução”, Atheneu: São Paulo. 2003.

SUBASIA, A. & ERCELEBI, E. Classification off eeg signals using neural network and logistic regression. *Computer Methods and Programs in Biomedicine*, vol. 78, no. , May 2005.

TAVARES, M. C. Monitor de Atividade Cerebral Baseado em Microcontrolador. Tese de Mestrado. Universidade Federal de Santa Catarina. Florianópolis. Agosto de 1997.

UNIVERSIDADE DE BONN. Departamento de Epileptologia da Universidade de Bonn – Alemanha, banco de dados de EEG. Disponível em: http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3. Acessado em: 02 de fevereiro 2010.

VOSS, A. e BAUMERT, M. Joint Symbolic Dynamics Analysis of Heart Rate and Systolic Blood Pressure Interactions in Dilated Cardiomyopathy. In: Proceedings of the 23rd Annual Engineering in Medicine and Biology Society Conference, pp. 555-558, Istambul. 2001.

VOSS, A., KURTHS, J., KLEINER, H. J., WITT, A., WESSEL, N., SAPARIN, P., OSTERZIEL, K. J., SCHURATH, R. e DIETZ, R. The Application of Methods of Non-Linear Dynamics for the Improved and Predictive Recognition of Patients Threatened By Sudden Cardiac Death, Cardiovascular Research, v. 31, n. 3 (Mar), pp. 419-433. 1996.

WINOGRAD, S. On computing the discrete fourier transform. Math. Computation, v. 32, 1978.

ZWE-LEE, G. Wavelet-based neural network for power disturbance recognition and classification. Power delivery, vol. 19, no. 4, Oct. 2004.