UNIVERSIDADE FEDERAL DE ALFENAS
DEPARTAMENTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

João Antonio da Silva

RECONHECIMENTO DE PADRÕES VIA REDES NEURAIS ART

Alfenas, 03 de Abril de 2013.

UNIVERSIDADE FEDERAL DE ALFENAS
DEPARTAMENTO DE CIÊNCIAS EXATAS
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

RECONHECIMENTO DE PADRÕES VIA REDES NEURAIS ART

João Antonio da Silva

Monografia apresentada ao Curso de Bacharelado em Ciência da Computação da Universidade Federal de Alfenas como requisito parcial para obtenção do Título de Bacharel em Ciência da Computação.

Orientadora: Profa. Dra. Angela Leite Moreno

Alfenas, 03 de Abril de 2013.

João Antonio da Silva

RECONHECIMENTO DE PADRÕES VIA REDES NEURAIS ART

A Banca examinadora abaixo-assinada aprova a monografia apresentada como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação pela Universidade Federal de Alfenas.

Prof. Mateus Neves Barreto

Universidade Federal de Alfenas

Prof. Ricardo Menezes Salgado

Universidade Federal de Alfenas

Profa. Angela Leite Moreno (orientadora)

Universidade Federal de Alfenas

Alfenas, 03 de Abril de 2013.

À minha mãe Eliana e aos meus irmãos Ana Paula e
Pedro Henrique, que por toda a vida
me amaram e apoiaram incondicionalmente.

AGRADECIMENTO

Este trabalho foi concluído com a ajuda que tive de diversas pessoas às quais serei eternamente grato.

Agradeço a todos os meus amigos que durante esses anos de faculdade, fizeram parte de minha vida, nos estudos e fora deles. Dentre as muitas amizades construídas nestes anos, devo mencionar: Renata Nobre, Diego Venturelli, Gabriela Pereira, Michele Sacramento e Elisabeth Barros. Obrigado por acreditarem em mim até quando eu mesmo não acreditava.

Agradeço aos meus professores, que contribuíram da melhor forma possível tanto para minha formação profissional, quanto para meu crescimento pessoal.

Meus agradecimentos aos professores Mateus Neves Barreto e Ricardo Menezes Salgado pela disponibilidade em participar da banca examinadora deste trabalho, pelas sugestões e contribuições que colaboraram para esta última versão do trabalho.

Agradeço à Professora Angela Leite Moreno, orientadora deste trabalho, pelo profissionalismo, comprometimento e disponibilidade. Mais que uma orientadora, uma amiga.

Agradeço a meus irmãos Ana Paula e Pedro Henrique, pelo apoio e compreensão durante todo este processo, vocês são o motivo pelo qual estou concluindo o curso.

Agradeço a minha mãe Eliana, pela pessoa maravilhosa que é, e por todo amor e dedicação em todos os momentos de minha vida, sobretudo nos últimos anos. Se hoje sou alguém devo isso a você, minha mãe, minha heroína.

Enfim, meu agradecimento a Deus, que colocou todas estas pessoas em minha vida.

“Em vez de pensar "quero fazer", pense "consigo fazer".
Em vez de pensar "quero ser", pense "posso ser".
Pensar "quero" é força da vontade,
mas pensar "posso" é força da convicção."

Katsumi Tokuhisa

RESUMO

As técnicas utilizadas para Reconhecimento de Padrões são aplicadas a um grande número de áreas científicas, como astronomia, cibernética, reconhecimento de caracteres, diagnósticos clínicos, análises sísmicas, avaliação de solos e produtos, dentre diversas outras. Nos últimos anos, tem se trabalhado muito para que o Reconhecimento de Padrões seja feito de forma automática com o uso de computadores da forma mais rápida e eficiente possível. Entretanto, o que se observa é o desenvolvimento de softwares que resolvem um problema específico, já que na maioria das vezes os padrões precisam passar por um pré-processamento antes de ser submetido ao processo de reconhecimento propriamente dito. Neste trabalho é apresentado um software de fácil utilização para se realizar o Reconhecimento de Padrões de qualquer base de dados sem se preocupar em realizar algum tratamento nos mesmos. O software utiliza para realizar a classificação dos padrões redes neurais artificiais baseadas na Teoria da Ressonância Adaptativa, pertencentes às famílias ART e ARTMAP. Para se utilizar o aplicativo, o usuário não precisa conhecer as características do classificador utilizado, apenas carregar os arquivos de entrada. Neste trabalho de conclusão de curso ainda é apresentado o resultado do reconhecimento de padrões utilizando o software desenvolvido para os seguintes problemas: Diagnóstico de câncer de mama, diagnóstico da doença de Parkinson, diagnóstico de doenças da coluna vertebral e identificação de movimentos da língua brasileira de sinais.

Palavras-Chave: Teoria da Ressonância Adaptativa, Reconhecimento de Padrões, Redes Neurais Artificiais.

ABSTRACT

The techniques used for pattern recognition are applied to a large number of scientific areas such as astronomy, cybernetics, character recognition, clinical diagnostics, seismic analysis, evaluation of soils and products, among many others. In recent years, has worked very for that Pattern Recognition is done automatically with the use of computers as fast and efficiently as possible. However, what is observed is the development of software that solves a specific problem, since in most cases the standards required to undergo a pre-processing before being subjected to the process of recognition itself. In this paper we present a user friendly software to perform the Pattern Recognition of any database without bothering to perform some of the same treatment. The software used to perform the pattern classification neural networks based on Adaptive Resonance Theory, belonging to the families ART and ARTMAP. To use the application, the user need not know the characteristics of the classifier used, just load the input files. In this study completion of course also presents the result of pattern recognition using software developed for the following problems: Diagnosing breast cancer, diagnosis of Parkinson's disease, diagnosis of spinal diseases and identification of tongue movements Brazilian signals.

Keywords: *Adaptative Ressonance Theory, Pattern Recognition, Artificial Neural Networks.*

LISTA DE FIGURAS

FIGURA 1 - ETAPAS ENVOLVIDAS EM UM SISTEMA DE RECONHECIMENTO DE PADRÕES. 32
FIGURA 2 - NEURÔNIO BIOLÓGICO 43
FIGURA 3 - NEURÔNIO ARTIFICIAL 44
FIGURA 4 – EXEMPLOS DE FUNÇÕES DE ATIVAÇÃO. 45
FIGURA 5 – PRINCIPAIS ARQUITETURAS DE REDES: (A) REDE NEURAL NÃO-RECORRENTE; (B) REDE NEURAL RECORRENTE. 46
FIGURA 6 – PRINCIPAIS MÉTODOS DE TREINAMENTOS: (A) TREINAMENTO SUPERVISIONADO; (B) TREINAMENTO NÃO-SUPERVISIONADO. 48
FIGURA 7 - ARQUITETURA BÁSICA DO SISTEMA ART. 55
FIGURA 8 – ARQUITETURA BÁSICA DO SISTEMA ARTMAP. 60
FIGURA 9 - ESTRUTURA DO SOFTWARE PROPOSTO. 63
FIGURA 10 – EXEMPLO DE ARQUIVO DE ENTRADA. 64
FIGURA 11 – TELA PRINCIPAL DO APLICATIVO. 66
FIGURA 12 – CONFIGURAÇÃO DA BUSCA EXAUSTIVA. 68
FIGURA 13 – GRÁFICO COMPARATIVO PARA OS RESULTADOS OBTIDOS COM A BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER*. 72
FIGURA 14 – GRÁFICO COMPARATIVO PARA OS RESULTADOS OBTIDOS PARA A BASE DE DADOS *VERTEBRAL COLUMN DATA SET*. 75
FIGURA 15 –GRÁFICO COMPARATIVO PARA OS RESULTADOS OBTIDOS PARA A BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET*. 78
FIGURA 16 - GRÁFICO COMPARATIVO PARA OS RESULTADOS OBTIDOS COM A BASE DE DADOS *MOVEMENT LIBRAS*. 80

LISTA DE TABELAS

TABELA 1 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 69
TABELA 2 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* MODIFICADA COM TREINAMENTO CONTINUADO REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 69
TABELA 3 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART EUCLIDIANA MODIFICADA REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 70
TABELA 4 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP *FUZZY* REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 70
TABELA 5 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP *FUZZY* MODIFICADA REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 70
TABELA 6 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP EUCLIDIANA MODIFICADA REFERENTE À BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 71
TABELA 7 – RESULTADOS OBTIDOS PARA A BASE DE DADOS *WISCONSIN DIAGNOSTIC BREAST CANCER.* 71
TABELA 8 MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* REFERENTE À BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 72
TABELA 9 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* MODIFICADA COM TREINAMENTO CONTINUADO REFERENTE À BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* .. 73
TABELA 10 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART EUCLIDIANA MODIFICADA REFERENTE *À* BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 73
TABELA 11 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP *FUZZY* REFERENTE À BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 73
TABELA 12 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP *FUZZY* MODIFICADA REFERENTE À BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 74
TABELA 13 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ARTMAP EUCLIDIANA MODIFICADA REFERENTE À BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 74
TABELA 14 - RESULTADOS OBTIDOS PARA A BASE DE DADOS *VERTEBRAL COLUMN DATA SET.* 75
TABELA 15 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 76
TABELA 16 - MATRIZ DE CONFUSÃO DA ANÁLISE OBTIDA PELA REDE ART *FUZZY* MODIFICADA COM TREINAMENTO CONTINUADO REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 76
TABELA 17 – MATRIZ DE CONFUSÃO OBTIDA PELA REDE ART EUCLIDIANA MODIFICADA REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 76
TABELA 18 - MATRIZ DE CONFUSÃO OBTIDA PELA REDE ARTMAP *FUZZY* REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 77
TABELA 19 - MATRIZ DE CONFUSÃO OBTIDA PELA REDE ARTMAP *FUZZY* MODIFICADA REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 77
TABELA 20 - MATRIZ DE CONFUSÃO OBTIDA PELA REDE ARTMAP EUCLIDIANA MODIFICADA REFERENTE À BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 77
TABELA 21 - RESULTADOS OBTIDOS PARA A BASE DE DADOS *OXFORD PARKINSON'S DISEASE DETECTION DATASET.* 78
TABELA 22 – RESULTADOS OBTIDOS COM A BASE DE DADOS *MOVEMENT LIBRAS.* 79

LISTA DE ABREVIAÇÕES

ART	*Adaptative Ressonance Theory*
DP	Doença de Parkinson
E	Espondilolistese
GN	*Grossberg Network*
H	Hérnia de Disco
HS	*Hybrid Search.*
IS	*Incremental Search*
KNN	*k-Nearest Neighbors*
MCS	*Monte Carlo Search*
MLP	*Multi-Layer Perceptron*
Modif.	Modificada
MTC	Modificada com Treinamento Continuado
N	Normal
RBF	*Radial Basis Function*
RNA	Redes Neurais Artificiais
SVM	*Support Vector Machine*
UCI	*University of California - Irvine*
VLSI	*Very Large Scale Integrtion*

SUMÁRIO

1 INTRODUÇÃO 25

1.1 JUSTIFICATIVA E MOTIVAÇÃO 27
1.2 PROBLEMATIZAÇÃO 27
1.3 OBJETIVOS 27
1.3.1 Gerais 27
1.3.2 Específicos 28
1.4 ORGANIZAÇÃO DA MONOGRAFIA 28

2 RECONHECIMENTO DE PADRÕES 31

2.1 CONSIDERAÇÕES INICIAIS 31
2.2 FASES DO RECONHECIMENTO DE PADRÕES 32
2.2.1 Template Matching 34
2.2.2 Abordagem Estatística 34
2.2.3 Abordagem Sintática ou Linguística 35
2.2.4 Abordagens Baseadas na Lógica *Fuzzy* 36

3 REDES NEURAIS ARTIFICIAIS 39

3.1 HISTÓRICO 39
3.2 CONSIDERAÇÕES GERAIS 41
3.2.1 Modelo Neural Biológico 42
3.2.2 Modelo Neural Artificial 43
3.3 ARQUITETURAS DE REDES NEURAIS ARTIFICIAIS 45
3.4 TREINAMENTO DE REDES NEURAIS ARTIFICIAIS 47
3.5 BENEFÍCIOS DAS REDES NEURAIS ARTIFICIAIS 49
3.6 APLICAÇÕES 51

4 TEORIA DA RESSONÂNCIA ADAPTATIVA (ART) 53

4.1 HISTÓRICO 53
4.2 CARACTERÍSTICAS BÁSICAS DO SISTEMA ART 55
4.2.1 Parâmetros de uma Rede ART 57
4.2.2 Alguns modelos neurais ART: 58
4.3 CARACTERÍSTICAS BÁSICAS DO SISTEMA ARTMAP 60
4.3.1 Parâmetros de uma Rede ARTMAP 61
4.3.2 Alguns modelos ARTMAP: 61

5 SOFTWARE E RESULTADOS 63

5.1 SOFTWARE 63
5.1.1 Base de dados 63
5.1.2 Núcleo 65
5.1.3 Interface 65
5.2 EXPERIMENTOS 68
5.2.1 Wisconsin Diagnostic Breast Cancer 69
5.2.2 Vertebral Column Data Set 72
5.2.3 Oxford Parkinson's Disease Detection Dataset 75
5.2.4 Movement Libras 79

6 CONCLUSÕES 81

6.1 CONCLUSÕES 81

6.2 SUGESTÕES PARA TRABALHOS FUTUROS 81
REFERÊNCIAS BIBLIOGRÁFICAS..... 83

1
Introdução

Este capítulo apresenta alguns detalhes sobre o trabalho proposto. Primeiramente apresenta-se uma visão geral deste trabalho, para, a seguir, apresentar-se a justificativa e motivação para o desenvolvimento do trabalho. Depois são apresentados os objetivos gerais e específicos deste trabalho. Enfim, apresenta-se a organização deste trabalho.

Reconhecimento de padrões é uma área de pesquisa cujo objetivo é classificar objetos em categorias ou classes. Neste contexto, diversos problemas abordando reconhecimento de padrões podem ser encontrados, desde o reconhecimento de imagem, movimento, voz e escrita, biometria, inclusive diagnóstico médico. A maioria desses problemas são resolvidos eficientemente pelos seres humanos, entretanto estas são atividades complexas para serem realizadas por computadores. Tais tarefas estão em fase avançada de desenvolvimento, mas ainda há muito o que ser feito para que o comportamento das máquinas se assemelhe ao observado nos seres humanos.

Nos problemas relacionados ao reconhecimento de padrões, nos quais as únicas informações conhecidas são os dados relativos aos problemas (dados de entrada), sem a existência de um ponto crítico para orientar quanto ao comportamento das soluções, é interessante um modelo matemático, mais especificamente de inteligência artificial, que classificasse esses dados explorando apenas as redundâncias contidas neles (Tauritz, 2005).

Deste modo, a aplicação de técnicas que não utilizam a supervisão do erro das saídas para a obtenção das classes, técnicas que explorem unicamente os elementos e as relações entre eles, gerando matematicamente estruturas de relacionamento entre os vários fatores que coexistem num mesmo cenário, são as mais interessantes.

Bishop (1995) assegura que a forma mais adequada para tratar problemas de reconhecimento de padrões é a abordagem estatística, entretanto ele considera as redes neurais artificiais como uma extensão dos métodos estatísticos. Contudo,

outras técnicas como *Template Matching*, abordagem sintática, métodos baseados na lógica *Fuzzy,* redes neurais artificiais e muitas outras são técnicas utilizadas para se realizar o reconhecimento de padrões da melhor maneira possível (Jain et. al., 2000).

Contudo, neste trabalho é proposto a aplicação de modelos neurais baseados na Teoria da Ressonância Adaptativa (*Adaptative Ressonance Theory* – ART). Essas redes englobam uma variedade ampla de redes neurais baseadas explicitamente na fisiologia e são definidas em relação ao seu algoritmo de treinamento.

Existem redes ART dos mais diversos tipos, utilizando paradigmas de aprendizado tanto supervisionado quanto não-supervisionado. Os modelos mais conhecidos e utilizados são:

- ART-1: Utiliza treinamento não-supervisionado. Foram os primeiros modelos desenvolvidos. Sua principal pecularidade é a utilização de dados binários como entrada da rede (CARPENTER, 1987a);
- ART-2: Utiliza treinamento não-supervisionado. Modelo desenvolvido para processar dados contínuos podendo também ser utilizada para entrada de dados binários (CARPENTER, 1987b);
- ART-3: Utiliza treinamento não-supervisionado. Modelo que utiliza a ação de neurotransmissores na propagação dos dados pela rede (CARPENTER, 1990);
- ARTMAP: Utiliza treinamento supervisionado. Composto por duas sub-redes: ART_a, que recebe e processa um determinado padrão de entradas apresentado, e ART_b, que constitui a resposta desejada para o padrão apresentado à rede (CARPENTER, 1991b);
- *Fuzzy* ART: Utiliza treinamento não-supervisionado. Este tipo de rede ART processa dados *Fuzzy*, *i.e.*, sua arquitetura realiza cálculos baseados na lógica nebulosa (CARPENTER, 1991a);
- *Fuzzy* ARTMAP: Utiliza treinamento supervisionado. Esta rede ARTMAP processa dados *Fuzzy* (CARPENTER, 1992).

1.1 Justificativa e Motivação

O problema de Reconhecimento de Padrões é um tema muito relevante atualmente, vários trabalhos, dentre eles podemos citar: Monitoramento de Sinais Via Satélite – GPS (Negri, 2012), Narizes Artificiais (Zanchettin e Ludermir, 2005), Classificação de Eventos, Grau de Torra de Café (Ribeiro, *et al, 2006*). Em cada um destes trabalhos é desenvolvido um software para a solução de problemas específicos. Assim, resolveu-se pela construção de um software para reconhecimento de padrões que possa ser utilizado para uma base de dados qualquer.

A principal motivação deste trabalho foi desenvolver um software multiplataforma para reconhecimento de padrões em tempo real em que o usuário não necessite ter conhecimento prévio sobre redes neurais artificiais.

1.2 Problematização

A tarefa de reconhecimento de padrões é usualmente trabalhosa e demorada. Os softwares existentes para tal tarefa são direcionados a uma única aplicação e, em geral, utilizam apenas uma rede neural artificial para a classificação. Então, como realizar o reconhecimento de padrões de um banco de dados qualquer utilizando mais de um tipo de rede sem grande custo computacional?

1.3 Objetivos

1.3.1 Gerais

O foco principal deste trabalho foi o desenvolvimento de um aplicativo de fácil utilização para reconhecimento de padrões utilizando redes neurais baseadas na Teoria da Ressonância Adaptativa.

1.3.2 Específicos

Os objetivos específicos desta monografia são:

- Implementar as redes ART:
    - *Fuzzy*
    - *Fuzzy* Modificada
    - *Fuzzy* Modificada com Treinamento Continuado
    - Euclidiana Modificada
    - Euclidiana Modificada com Treinamento Continuado
- Implementar as redes ARTMAP:
    - *Fuzzy*
    - *Fuzzy* Modificada
    - Euclidiana Modificada
- Desenvolver um software simples e de fácil utilização para reconhecimento de padrões:
    - Escolha da rede neural pelo usuário;
    - Realização de busca exaustiva para obtenção do melhor conjunto de parâmetros a ser utilizado pela rede neural;
    - Visualização do resultado pelo usuário.

Apresentar os resultados obtidos com a utilização de quatro bases de dados para teste das redes neurais.

1.4 Organização da Monografia

O presente trabalho de conclusão de curso está organizado da seguinte forma:

- **Reconhecimento de Padrões:** Este capítulo é voltado a introduzir o problema a ser abordado pela rede neural: Reconhecimento de Padrões.

Nele são apresentados os principais conceitos envolvendo reconhecimento de padrões e uma descrição rápida das principais técnicas utilizadas para sua resolução.

- **Redes Neurais Artificiais:** Neste capítulo, apresentam-se os principais conceitos de redes neurais, inclusive uma visão histórica de sua evolução.
- **Teoria da Ressonância Adaptativa:** Este capítulo é dedicado à Teoria da Ressonância Adaptativa. Apresenta-se primeiramente a revisão história para, a seguir, a descrição dos modelos utilizados neste trabalho.
- **Aplicativo e Resultados:** Neste, apresenta-se a interface do aplicativo bem como a estrutura do arquivo de entrada. A seguir, são apresentados os resultados obtidos com a utilização do aplicativo, as matrizes de confusão e uma tabela comparativa dos resultados da literatura e com os obtidos neste trabalho.
- **Conclusão:** O objetivo deste capítulo é apresentar as conclusões gerais obtidas neste trabalho de conclusão de curso bem como as sugestões de trabalhos futuros.

2
Reconhecimento de Padrões

Neste capítulo são abordados alguns conceitos sobre o problema de reconhecimento de padrões. Na seção 2.1 são apresentados alguns conceitos que permitem o leitor se familiarizar com o problema. Em seguida, é apresentada uma síntese com as fases que compõe o desenvolvimento de um sistema de reconhecimento de padrões, com maior detalhamento nas abordagens utilizadas para se realizar o reconhecimento ou classificação de padrões propriamente dita. Este capítulo se encerra com a apresentação de áreas nas quais as técnicas de reconhecimento de padrões podem ser aplicadas.

2.1 Considerações Iniciais

O termo reconhecimento de padrões foi introduzido na década de 1960, e originalmente consistia na detecção de formas simples. É uma habilidade extremamente desenvolvida nos seres humanos e em alguns animais. A facilidade que o ser humano possui em reconhecer padrões pode ser atribuída aos estímulos aos quais o indivíduo foi submetido anteriormente, de forma que o reconhecimento de padrões agrega algum aprendizado e evolui com a experiência. Contudo esta é uma tarefa bastante complexa para as máquinas, e é um grande desafio desenvolvê-las de forma que apresentem tais características (Castro e Prado, 2002).

Segundo Duda e Hart (1974), reconhecimento de padrões é o "campo interessado no reconhecimento de regularidades significativas em ambientes ruidosos ou complexos, feito por máquinas" ou a "procura por uma estrutura de dados". De forma mais simples, podemos considerar que reconhecimento de padrões é a área da pesquisa que busca identificar certas estruturas nos dados de entrada em comparação a estruturas conhecidas e classificar objetos (padrões) em um número de categorias ou classes. É interessante ressaltar que o grau de associação é maior entre padrões de uma mesma classe e menor entre padrões de classes diferentes.

Gonzales e Tou (1981) definem padrão como "uma descrição quantitativa ou estrutural de um objeto", de forma que as propriedades de um dado padrão possibilitem o agrupamento de objetos com características semelhantes dentro de uma classe ou categoria. Uma classe é um agrupamento de padrões que compartilham propriedades comuns.

2.2 Fases do Reconhecimento de Padrões

Duda e Hart (1974) consideram que os sistemas de reconhecimento de padrões geralmente seguem uma sequência de etapas bem definidas, sendo que a complexidade de cada uma dessas etapas está diretamente relacionada com a natureza do problema em questão. As etapas atestadas por Duda e Hart podem ser observadas na Figura 1.

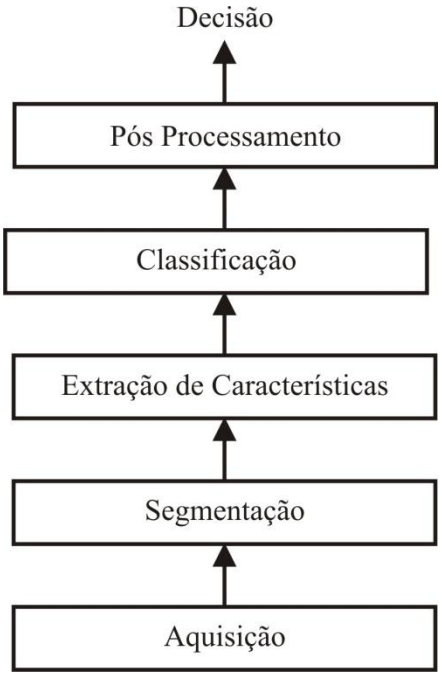


Figura 1 - Etapas envolvidas em um sistema de reconhecimento de padrões.

A primeira etapa consiste na aquisição dos dados de entrada para o sistema e frequentemente é caracterizada pelo pré-processamento e normalização dos dados.

Na segunda etapa é feita a segmentação dos dados, ou seja, esta etapa é responsável por separar os objetos que são interessantes para o reconhecimento daqueles que não o são. Duda e Hart afirmam que a segmentação é um dos problemas mais complexos no desenvolvimento de sistemas para o reconhecimento de padrões.

A próxima fase é denominada extração de características e é responsável pela formação de um conjunto de medidas que apresentem alguma similaridade. Ressalta-se que um bom sistema de extração de características pode facilitar o trabalho do reconhecedor, uma vez que a similaridade entre objetos de uma mesma classe são maiores do que a similaridade entre objetos de classes distintas (Moreno, 2010). Assim, recomenda-se que seja feita uma seleção adequada das características que irão descrever um padrão. Um conjunto ideal de características é aquele que é composto pelo menor número de atributos de tal maneira que não prejudique o desempenho do classificador (Duda e Hart, 1974).

A quarta etapa é o reconhecimento propriamente dito. Nela, os padrões de entrada são submetidos ao classificador, que deve atribuir uma classe a cada padrão de entrada. A classificação pode ser supervisionada, quando o padrão de entrada é identificado em relação às classes predefinidas pelos padrões de treinamento. Pode ser também não supervisionada, quando o padrão é associado a uma classe que é aprendida em função da similaridade dos padrões de treinamento. Muitos problemas podem surgir durante esta fase, mas o principal deles é encontrar um reconhecedor adequado para um dado problema, de forma a maximizar o seu desempenho. O treinamento supervisionado e o não supervisionado são detalhados no próximo capítulo.

A quinta e última fase pode ou não compor um sistema de reconhecimento de padrões. Envolve a análise dos resultados obtidos pelo reconhecedor e a tomada de uma decisão.

O foco deste trabalho está na quarta etapa, isto é, no desenvolvimento de um reconhecedor de padrões. Além disso, será analisado os resultados obtidos para alguns problemas, cujos conjuntos de padrões a serem analisados já passaram pelas primeiras etapas no sistema e não entraremos no mérito neste momento.

Muitas são as técnicas para os problemas envolvidos em reconhecimento de padrões. (Jain et. al., 2000) classifica essas abordagens em: *Template Matching*,

abordagem sintática, métodos baseados na Lógica *Fuzzy*, abordagem estatística e redes neurais artificiais. Alguns trabalhos, como o apresentado por (BISHOP, 1995), consideram a abordagem de Reconhecimento de Padrões utilizando redes neurais como uma extensão da abordagem estatística convencional, sendo a segunda forma considerada como uma evolução da primeira. Pode-se ainda citar os seguintes métodos de classificação (Jain et. al., 2000): Discriminante linear de Fisher, Árvores de decisão binária, Rede de base radiais, classificador logístico, a Regra dos K-vizinhos mais próximos e as Máquinas de Vetor Suporte.

A técnica compreendida neste trabalho será redes neurais artificiais, mais precisamente nas redes neurais baseadas na Teoria da Ressonância Adaptativa, que serão mais bem detalhadas nos próximos capítulos. A seguir será feita uma descrição sucinta de alguns dos métodos mais conhecidos, indicando os diferentes tipos de problemas e de aplicações envolvidas em cada uma dessas abordagens.

2.2.1 Template Matching

Template Matching (Theodoridis, 1999), a mais antiga das abordagens em reconhecimento de padrões, é indicada para os problemas em que os padrões a serem reconhecidos são bastante próximos do modelo padrão.

Consiste em modelos de protótipos que contém os padrões a serem reconhecidos e cujas características servem de parâmetro para a comparação. Os padrões de entradas são submetidos a uma operação genérica e então é calculada uma medida de similaridade entre as entradas e os protótipos. A medida de similaridade, que geralmente está relacionada à ideia de correlação, deve ser otimizada em função do conjunto de treinamento. A maior desvantagem desta abordagem reside no alto custo computacional. Ainda pode-se obter classificações inadequadas se os padrões apresentarem ruídos.

2.2.2 Abordagem Estatística

A abordagem estatística reconhece a natureza probabilística dos padrões apresentados ao reconhecedor e a forma como os resultados devem ser expressos. Assim, nesta abordagem, o objetivo é minimizar a probabilidade de erro durante o

reconhecimento dos padrões. Para Schalkoff (1992), o reconhecimento de padrões estatístico utiliza um conjunto de medidas estatísticas que denotam as características extraídas dos dados de entrada, os quais são utilizados para associar cada vetor de características a uma classe.

A fronteira de decisão é determinada para cada classe na teoria da decisão estatística. Esta fronteira pode ser especificada e, neste caso, observamos o aprendizado supervisionado onde há a presença de um especialista que produz as categorias em um conjunto de treinamento. Já no caso do aprendizado não-supervisionado, não existe um especialista e o sistema gera agrupamentos naturais dos padrões de entrada, sendo este agrupamento definido explicita ou implicitamente pela própria classe (Duda et. al., 2001). Ressalta-se que no aprendizado não-supervisionado, algoritmos diferentes de classificação produzirão resultados diferentes.

Os métodos de reconhecimento estatístico ainda podem ser classificados em paramétricos e não-paramétricos. Na abordagem paramétrica é exigido um determinado conhecimento à priori sobre a distribuição probabilística que melhor se ajusta aos dados. Quando não se tem conhecimento sobre a distribuição dos dados e não for possível supor um bom modelo que os represente, é aconselhável utilizar métodos não-paramétricos, já que estas são robustas e não requerem informações da verdadeira distribuição dos dados (Jain et. al., 2000).

2.2.3 Abordagem Sintática ou Linguística

Os padrões são abordados sob uma perspectiva hierárquica, isto é, um padrão é composto por subpadrões, que por sua vez também são compostos por subpadrões. Os subpadrões elementares são denominados primitivas e um padrão complexo é caracterizado pelas relações entre estas primitivas. Este comportamento pode ser observado na sintaxe de uma linguagem, em que o alfabeto são as primitivas, e os padrões são sentenças pertencentes à linguagem, produzidas a partir da respectiva gramática. A gramática para cada classe de padrões deve ser inferida a partir dos exemplos de treinamento.

Uma vantagem potencial da abordagem sintática é que, além da classificação, ela também permite uma descrição de como os padrões são gerados a

partir das primitivas. É útil em problemas em que os padrões não podem ser bem descritos em medidas numéricas, ou naqueles nos quais os padrões são tão complexos que suas especificidades não podem ser enunciadas. A abordagem sintática é bastante utilizada em problemas de reconhecimento de voz, mas também em reconhecimento de caracteres e eventos temporais em vídeos.

A desvantagem deste método está na grande demanda computacional. Isto porque inicialmente os padrões devem ser segmentados a fim de se obter as primitivas e na inferência da gramática, o que leva a um grande número de elementos combinatórios, isto é, das possibilidades a serem investigadas.

2.2.4 Abordagens Baseadas na Lógica *Fuzzy*

A teoria dos conjuntos *Fuzzy (*ou nebulosa*)* foi proposta por (Zadeth, 1965), com inspiração no comportamento humano que interpreta o mundo sem precisão e na descrição desses por atributos linguísticos. A partir da lógica apresentada por Zadeth, é possível desenvolver técnicas de decisão objetiva baseada em critérios subjetivos. A partir de conceitos como "muito quente", "quente", "meio quente", "um pouco quente", "um pouco frio", "meio frio", "frio", "muito frio", podem ser definido critérios de decisão. Estes conceitos subjetivos característicos do estado físico ambiental do ser humano estão relacionados a capacidade que o mesmo possui em atribuir um valor qualitativo a certos padrões.

A classificação utilizando a lógica *Fuzzy* geralmente é feita utilizando regras nebulosas, ou seja, regras de inferência que utilizam delimitadores dos valores de pertinência de cada padrão aos conjuntos, e métodos de defuzzificação. Para um estudo mais detalhado dos conceitos da lógica *Fuzzy*, recomenda-se a leitura do Apêndice B do trabalho de Moreno (2010).

A maioria dos autores afirma que não há um classificador melhor que outro. O que se observa é o melhor desempenho de um classificador em detrimento de outros, o que varia em função das características de funcionamento do reconhecedor que são mais adequadas para o dado problema. Não é rara também a combinação de alguns desses métodos (Tou e Gonzales, 1981). A combinação de previsores é denominada *ensembles* (agrupamentos), e como sugere, consiste na

combinação dos resultados de vários previsores distintos em busca de um resultado melhor (ou mais estável) que os obtidos pelos componentes individuais.

Dezenas de áreas podem ser destacadas na aplicação de reconhecimento de padrões, dentre elas: diagnósticos médicos (Steiner, 1995), visão computacional (Perelmuter et. al., 1995), reconhecimento de voz (Ferreira, 1994), processamento de imagens, análise de caracteres manuscritos (Prado e Elfes, 1975), sinais de radar (Mascarenhas, 1987) e reconhecimento e identificação de cromossomos (Todesco, 1995). Vale lembrar que em algumas das aplicações de reconhecimento podem requerer uma ocorrência exata de um padrão, ao passo que outras são satisfeitas quando uma aproximação é encontrada.

3
Redes Neurais Artificiais

Este capítulo apresenta um resumo dos principais conceitos utilizados no contexto das redes neurais artificiais. Na primeira seção é feito um levantamento histórico dos principais marcos do desenvolvimento da área. A seguir são apresentados alguns conceitos fundamentais sobre as redes neurais, bem como uma breve analogia entre os neurônios biológicos e os neurônios artificiais. Ainda são apresentadas as principais arquiteturas de redes e paradigmas para treinamento das mesmas. O capítulo termina com os benefícios oferecidos pelas redes neurais na solução de problemas e algumas áreas nas quais elas são amplamente empregadas.

3.1 Histórico

O trabalho de Warren McCulloch e Walter Pitts publicado em 1943, é considerado como o marco inicial da área de redes neurais artificiais. Em seu artigo *A Logical Calculus of the Ideas Immanent in Neural Activity* (MacCulloc & Pitts, 1943) McCulloc e Pitts uniram estudos neurobiológicos com a lógica matemática e propuseram a criação de neurônios formais. Neste trabalho, eles interpretam o funcionamento do neurônio biológico como sendo um circuito binário simples que combina várias entradas e gera um sinal de saída, caracterizado pela soma ponderada dos valores de entrada. Em 1947, McCulloch e Pitts conectaram um número suficiente desses neurônios de forma a construir uma rede e mostraram que, com os devidos ajustes, essas redes seriam capazes de executar funções complexas.

Outro marco importante no desenvolvimento das redes neurais artificiais foi a introdução do conceito de aprendizagem das redes neurais. No livro *The Organization of Behavior* (Hebb, 1949), Hebb propõe um método de aprendizagem baseado no comportamento fisiológico do processo de aprendizagem que acontece no cérebro humano. Hebb considerou a hipótese de que, à medida que diferentes tarefas vão sendo aprendidas pelo organismo, as conexões internas do cérebro passam por contínuas alterações. Este processo faz com que a conexão entre dois

neurônios seja reforçada, na medida em que o primeiro neurônio ativa o segundo através da sinapse. O conceito de aprendizagem hebbiana é ainda é utilizado no desenvolvimento de redes neurais atualmente.

Em 1958, Frank Rosenblant propõe o modelo *perceptron*, baseado no modelo McCulloch-Pitts (Rosenblant, 1958). Este modelo caracteriza-se por possuir três camadas. A camada de entrada recebe os dados exteriores, gera impulsos e os transmite para a segunda camada através de conexões que podem ter sua eficiência de transmissão (peso) ajustada. A terceira camada, a camada de saída, recebe os estímulos da segunda camada e é responsável por apresentar a resposta.

No ano de 1960 Widrow e Hoff apresentaram o modelo *Adaline* (*ADAptative LINear Element*) (Widrow & Hoff, 1960). Este modelo se difere do *perceptron* quanto ao procedimento de aprendizagem. Dois anos mais tarde, Widrow e alguns estudantes propuseram a estrutura conhecida com *Madaline* (*Multiple-Adaline*) (Widrow, 1962). Durante esta década, acreditava-se que as redes neurais artificiais poderiam realizar qualquer tipo de trabalho.

No entanto, em 1969, Minsky e Papert apresentam a prova matemática de que a convergência do *perceptron* desenvolvido por Rosenblatt era limitada aos problemas linearmente separáveis. Assim, tarefas simples como o problema do ou – exclusivo, não podiam ser resolvidos por tal modelo. A publicação, que também recebeu o nome de *Perceptrons* (Minsky and Papert, 1969), desencorajou as agências de fomento a desenvolver novas pesquisas e consequentemente os pesquisadores na área.

Assim, a década de 1970 é considerada uma época de adormecimento para as redes neurais e poucas atividades foram desenvolvidas na área. O algoritmo de retropropagação (*backpropagation*) apresentado por Werbos (Werbos, 1974) merece destaque neste período. Segundo Marchiori (2006), o algoritmo desenvolvido por Werbos "possui a capacidade de treinar redes multicamadas" e viria a ser objeto de estudo algum tempo mais tarde.

O trabalho de John Hopfiel em 1982 (Hopfield, 1982) apresenta um modelo de rede que associa o estado da rede a uma função energia. Isso garante que a rede se estabiliza na ausência de entradas externas, isto é, a rede se estabiliza quando chegasse a um mínimo local ou global da função energia.

Ainda em 1982, Kohonen propôs um método de aprendizagem competitivo conhecido como auto-organizável (Kohonen, 1982). Neste modelo, estabeleceu-se um método de aprendizagem competitivo, não-supervisionado, que considera que os neurônios se conectam com aqueles posicionados mais próximos a ele.

Em 1986, Rumelhart, Histon e Williams redescobrem o trabalho proposto por Werbos em 1974 e utilizam o algoritmo de retropropagação em várias aplicações, consolidando este algoritmo como o algoritmo de aprendizagem mais usado no treinamento de redes *perceptrons* com múltiplas camadas. A notoriedade do algoritmo fez com que a comunidade científica voltasse a se interessar pelo campo das redes neurais artificiais. Desde então, os estudos na área tem se expandido notavelmente.

O modelo ART (*Adaptative Resonance Theory*) foi proposto por Gail Carpenter e Stephen Grosberg (Carpener & Grossberg, 1987a; Carpenter & Grossberg, 1987b) em 1987. Nesse modelo utiliza-se um mecanismo de treinamento não-supervisionado e competitivo. Nele, quando o padrão de entrada é compatível com um dos padrões aprendidos pela rede, ocorre um processo chamado de ressonância adaptativa (Moreno, 2010).

Atualmente, existem várias arquiteturas neurais propostas. E mesmo que, em teoria, as redes neurais artificiais possam ser utilizadas para resolver qualquer tipo de problema computacional, o que se observa é sua aplicação em problemas mais específicos. É interessante ressaltar que cada arquitetura de rede apresenta resultados melhores para determinados tipos de problemas e assim, para Kohonen (1989) a contribuição das redes neurais artificiais abrange diversas áreas do conhecimento, principalmente nas áreas de representação do conhecimento e reconhecimento de padrões.

3.2 Considerações Gerais

As Redes Neurais Artificiais (RNA) tem mostrado ser uma poderosa ferramenta na classificação e reconhecimento de padrões. RNAs são técnicas computacionais que representam um modelo matemático inspirado na estrutura neural de organismos

inteligentes e são capazes de aprender através da experiência e generalização adquiridos com dados passados (Ferreira, 2003; Salgado, 2004).

3.2.1 Modelo Neural Biológico

O neurônio é a unidade celular fundamental do sistema nervoso humano (Kartalopuolos, 1996). O cérebro é composto por cerca de 100 bilhões de neurônios que se conectam numa complexa rede de comunicação e que tem papel essencial na determinação do comportamento do corpo humano. Basicamente, um neurônio biológico tem a seguinte formação:

- **Corpo celular** (do grego, *sôma*): É o processador do neurônio, responsável pela recepção, processamento e geração dos impulsos nervosos.
- **Sinapse:** Ligação química e temporária é responsável pela transmissão dos impulsos nervosos entre os neurônios.
- **Dendrito:** Recebem os impulsos nervosos de outros neurônios e conduzem estes sinais de entrada até o corpo celular.
- **Axônio:** Comporta-se como um canal de saída do neurônio. É responsável por transmitir os sinais processados pelo corpo celular aos demais neurônios. Perceba que um axônio divide-se em vários ramos (dendritos), que se conectam com os demais neurônios de forma a estabelecer uma rede.

Isto pode ser verificado na Figura 2.

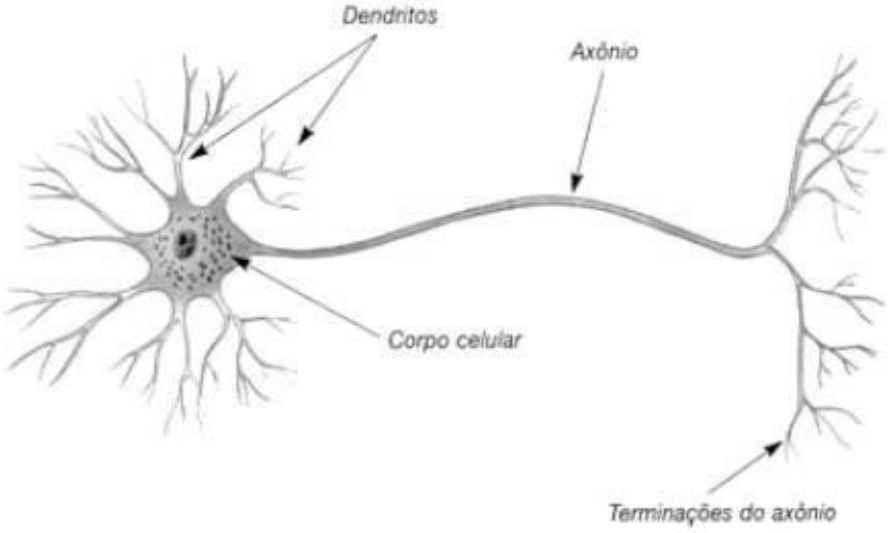


Figura 2 - Neurônio Biológico

3.2.2 Modelo Neural Artificial

O elemento fundamental de uma RNA é o neurônio (ou elemento processador) que tem por objetivo simular o comportamento de um neurônio biológico. Assim como o sistema nervoso dos humanos, uma RNA é composta por um conjunto de neurônios conectados de forma a constituir uma grande rede de comunicação, recebendo, armazenando e transmitindo informações.

O modelo apresentado por McCulloch-Pitts (McCulloch e Pitts, 1943), foi o primeiro modelo formulado (Figura 3). Neste modelo, a atividade u_k é a soma de m entradas $x_1, x_2, \ldots, x_p$ que representam as características dos padrões apresentados à rede. A cada uma dessas estradas, está associada um peso $w_{k\,j}$ que representam os tipos de sinapses utilizados pela rede. As sinapses podem ser inibitórias, quando a troca de informações entre os neurônios são impedidas ou dificultadas, ou excitatórias, quando a troca de informações entre os neurônios é permitida. Após a multiplicação das entradas x_j com os pesos $w_{k\,j}$, o corpo celular é responsável por realizar a soma dos valores $x_j w_{k\,j}$. Uma função de ativação não-linear $\varphi(.)$ é então aplicada a este somatório a fim de ativar ou não a saída do neurônio e enviar este sinal aos demais neurônios que constituem a rede.

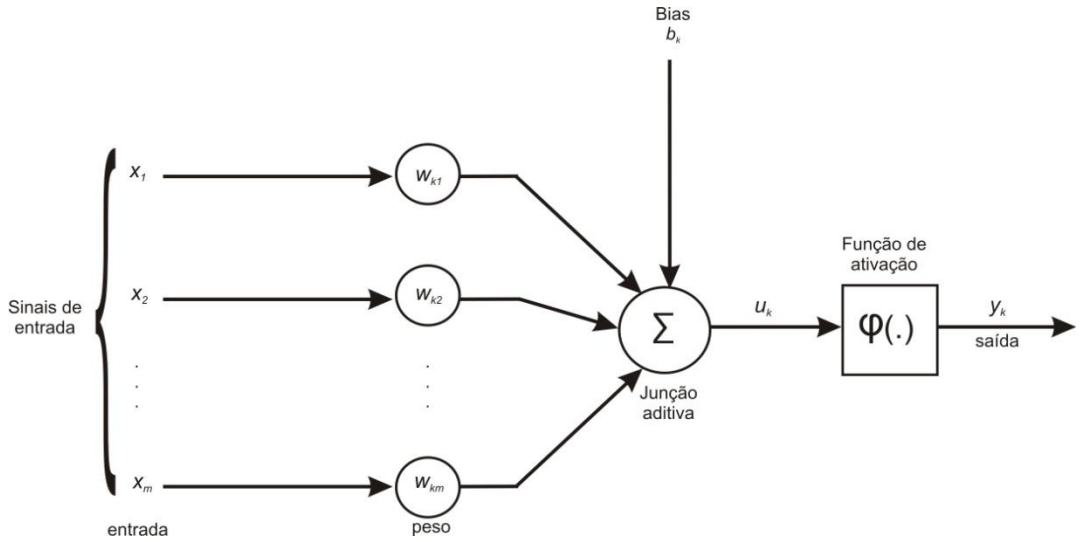


Figura 3 - Neurônio Artificial

Ressalta-se que a unidade proposta por McCulloch-Pitts é o mais popular modelo de neurônio artificial. Mas, assim como existem diversos neurônios biológicos especializados para algum tipo de processamento, existem especificações dos neurônios artificiais, a fim de adequá-los a um problema específico. As modificações mais comuns aplicadas ao neurônio de McCulloch-Pitts consistem no uso de outras funções de ativação (Mendes Filho & Carvalho, 1997). Várias funções de ativação podem ser utilizadas, dependendo do objetivo da rede. Na Figura 4 são apresentados exemplos das funções de ativação mais utilizadas.

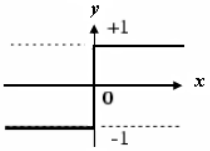
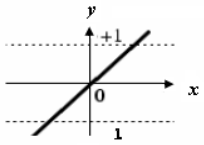
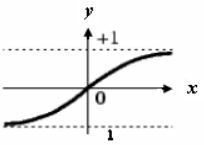
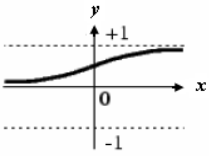
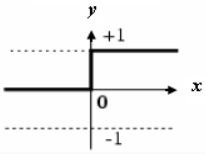
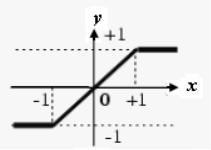
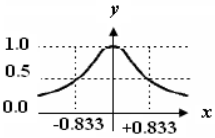
Função	Gráfico	Equação
Degrau Bipolar		Se $x > 0 \rightarrow y = 1$ Se $x = 0 \rightarrow y = 0$ Se $x < 0 \rightarrow y = -1$
Linear		$y = x$
Tangente Hiperbólica		$y = \frac{1 - e^{-\beta x}}{1 + e^{\beta x}}$
Logística (Sigmoide)		$y = \frac{1}{1 + e^{-\beta x}}$
Degrau		Se $x \geq 0 \rightarrow y = 1$ Se $x < 0 \rightarrow y = 0$
Rampa		Se $x > 1 \rightarrow y = 1$ Se $-1 \leq x \leq 1 \rightarrow y = x$ Se $x < 0 \rightarrow y = -1$
Base Radial		$y = e^{-x}$

Figura 4 – Exemplos de funções de ativação.

3.3 Arquiteturas de Redes Neurais Artificiais

Dá-se o nome de arquitetura de rede neural à maneira como os neurônios estão estruturados em uma rede. Quanto a esta estrutura, que se refere ao método de propagação da informação recebida, uma rede neural pode ser classificada em dois tipos (Figura 5):

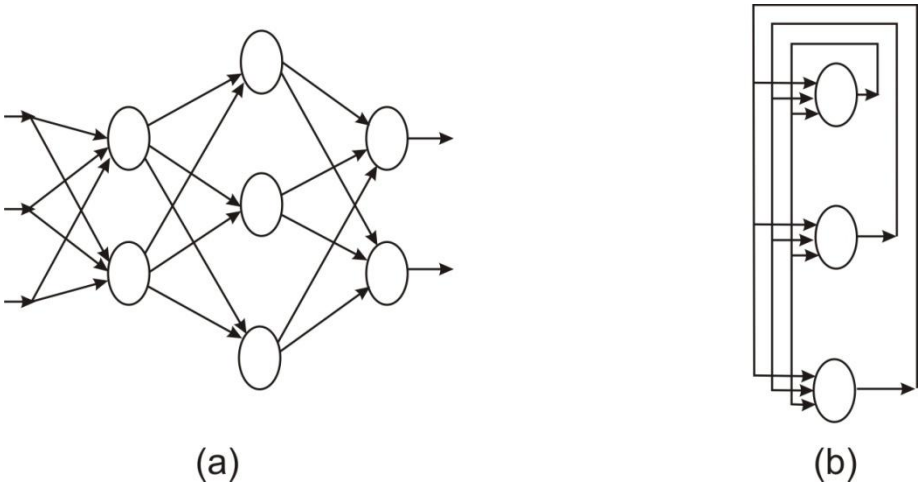


Figura 5 – Principais arquiteturas de redes: (a) Rede neural não-recorrente; (b) Rede neural recorrente.

- Redes Neurais Artificiais Não-Recorrentes: Segundo Isoda (2009), uma rede neural não-recorrente "caracteriza-se estruturalmente por estar disposta em camadas. Nestas redes, cada camada de neurônios recebe sinais somente das camadas anteriores, ou seja, elas não possuem laço de realimentação". As redes neurais não-recorrentes, também conhecidas como *feedforward*, podem ser formadas por uma ou várias camadas e não possuem memória. Ainda, a saída é exclusivamente determinada em função da entrada dos valores dos pesos. As principais redes neurais pertencentes a esta classe são: perceptron (Widrow and Lehr, 1990), as redes neurais da família ART (Carpener & Grossberg, 1987a; Carpenter & Grossberg, 1987b) e a rede Kohonen (Kohonen, 1989).

- **Redes Neurais Artificiais Recorrentes:** Uma rede é dita recorrente "se ela contém laços de realimentação, ou seja, contém conexões das saídas de uma determinada camada para a entrada da mesma ou de camadas anteriores" (Isoda, 2009). Uma rede neural recorrente se distingue da não-recorrente por ter pelo menos um laço de realimentação. E diferentemente das redes não-recorrentes, estas redes desenvolvem uma memória em longo prazo nos neurônios internos. As redes mais relevantes desta classe são as redes de Hopfield (Simpson, 1989).

3.4 Treinamento de Redes Neurais Artificiais

A propriedade mais interessante que as redes neurais artificiais herdaram do sistema nervoso humano é a habilidade de aprender e melhorar seu desempenho através do conhecimento que adquiriram durante o treinamento. (Ferreira, 2003). O treinamento de uma rede neural pode ser definido como um processo iterativo onde as sinapses são ajustadas de forma que generalizem e reproduzam o raciocínio aproximado. O treinamento de uma rede neural também pode ser definido como um processo de ajustes aplicados aos pesos de maneira que um conjunto de entrada produza o conjunto de saída desejada. A ideia é que a rede se torne mais instruída à medida que vai adquirindo conhecimento de seu ambiente, através do processo de treinamento.

Algoritmo de treinamento é um conjunto de regras predeterminadas para resolver um problema específico. Geralmente, os algoritmos de aprendizagem se diferem na forma em que a adaptação dos pesos sinápticos de um neurônio é realizada e pela maneira pela qual a rede neural interage com o seu ambiente. Dessa maneira, os treinamentos são classificados em:

- **Treinamento Supervisionado:** Neste paradigma de treinamento, o ajuste dos pesos sinápticos é realizado através da combinação dos padrões de entrada e saída. Há também a presença de um agente externo (professor ou tutor), que avalia o desempenho da rede e adquire conhecimento da mesma na forma de mapeamentos de entrada-saída. Para cada entrada apresentada à rede, é verificado se a saída obtida corresponde a saída desejada. Se a saída obtida não for a desejada, é calculado o erro (diferença entre a saída obtida e a desejada), e a partir deste deve-se atualizar os pesos sinápticos de forma que a rede armazene o conhecimento desejado. Essa iteratividade deverá ser repetida em todo o conjunto de treinamento, até que a saída obtida esteja próxima o suficiente da saída desejada (Lopes, 2005). Um dos algoritmos mais difundidos para este tipo de treinamento é o algoritmo de retropropagação (*backpropagation*).

- **Treinamento Não-Supervisionado:** Neste método, também conhecido como aprendizagem auto-organizada (*self-organization*), apenas o padrão de entrada é fornecido à rede. A rede processa as entradas que lhe são apresentadas e classifica-as conforme seus próprios critérios (Marchiori, 2002). Neste paradigma de aprendizado, a rede é capaz de extrair correlações dos padrões de que lhe foram apresentados, o que possibilita a criação de representações internas que representam características da entrada e criar novas classes sempre que necessário. De maneira mais simples, os neurônios atuam como classificadores enquanto que os dados de entrada são os elementos a serem classificados de acordo com as relações aprendidas durante o treinamento. As redes ART possuem treinamento não-supervisionado.
- **Treinamento Híbrido:** Consiste na combinação do treinamento supervisionado com o treinamento não-supervisionado. Neste paradigma, parte dos pesos sinápticos são atualizados através do treinamento supervisionado enquanto outros são ajustados através das correlações adquiridas pelo treinamento não-supervisionado.

A Figura 6 (a) ilustra o treinamento supervisionado. Já na Figura 6 (b) está representado o treinamento não-supervisionado.

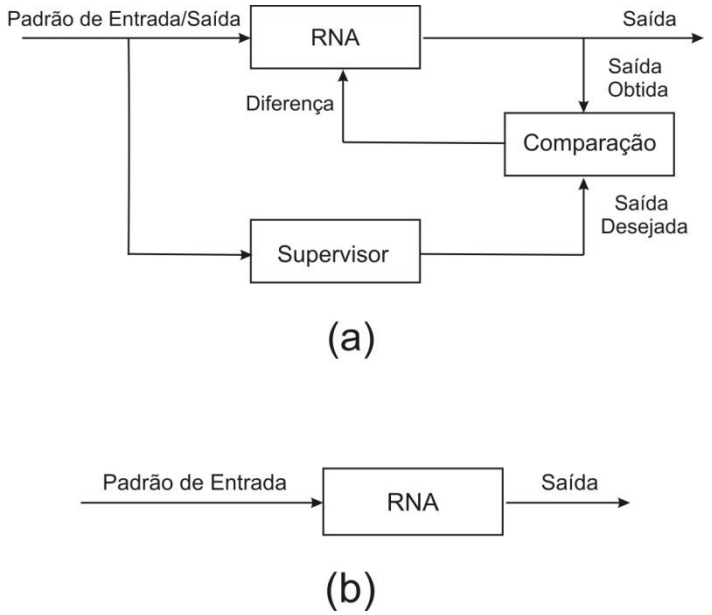


Figura 6 – Principais métodos de treinamentos: (a) Treinamento supervisionado; (b) Treinamento não-supervisionado.

Há ainda outras técnicas de treinamento mais específicas, dentre os quais merecem destaque os treinamentos por reforço, competitivo, regra delta, gradiente descendente e Hebbiano (Ferreira, 2003).

3.5 Benefícios das Redes Neurais Artificiais

Uma das vantagens das redes neurais artificiais é o poder de processamento altamente paralelo de sua estrutura. Mas é a habilidade de aprendizagem, e consequente generalização, que faz com que as redes neurais sejam uma classe muito especial de sistemas. A generalização se refere ao fato de a rede neural produzir saídas adequadas quando entradas que não estavam presentes no treinamento são apresentadas a rede. Estas duas características permitem que as redes neurais resolvam problemas altamente complexos que são normalmente difíceis de ser tratados.

É importante observar que as redes neurais não solucionam problemas sozinhas. Na prática, sua aplicação deve ser feita de forma criteriosa em conjunto com uma abordagem consistente de engenharia de sistemas. Assim, um problema complexo deve ser decomposto em tarefas mais simples, de maneira que se possa atribuir às redes neurais as tarefas que tem capacidade de resolver, explorando o máximo de habilidade de cada tipo de rede.

As redes neurais apresentam características singulares, e oferece propriedades interessantes aos sistemas que as utilizam:

- **Não-linearidade:** É uma propriedade muito importante dos neurônios, principalmente nos casos onde a entrada é inerentemente não-linear, o que é observado na maioria dos sistemas naturais.
- **Mapeamento de Entrada-Saída:** No aprendizado supervisionado, um dos paradigmas de aprendizado, os ajustes dos pesos sinápticos são realizados a partir de exemplos, desenvolvendo um mapeamento entre a entrada e a saída. Para cada padrão de entrada, é determinado o ajuste dos pesos que minimizam a diferença entre a resposta apresentada pela rede e a resposta desejada para a respectiva entrada.

- **Adaptabilidade:** As redes neurais tem a capacidade de adaptar seus pesos sinápticos em resposta às alterações do meio em que se encontra. Através de um retreinamento, uma rede neural pode facilmente lidar com pequenas alterações nas condições operativas do ambiente. Ainda, se a rede neural está inserida em um ambiente não-estacionário, ou seja, onde as estatísticas mudam com o tempo, a rede pode ser adaptada para ajustar seus pesos em tempo real. Contudo, esta adaptação deve ser feita com cuidado, já que a rede pode passar a responder às entradas às quais não deveria responder, o que compromete a robustez da rede. Este problema é conhecido com o dilema da estabilidade-plasticidade (Grossberg, 1988b).
- **Resposta a Evidências:** No contexto de classificação de padrões, uma rede neural pode ser projetada para fornecer informações sobre a confiança da decisão que foi tomada. Essas informações podem ser utilizadas para melhorar o desempenho de classificação da rede, a partir da rejeição de padrões ambíguos.
- **Informação Contextual:** Dependendo do método de treinamento adotado, a saída apresentada pela rede pode ser afetada por todos os neurônios que a constituem, e o ajuste dos pesos pode atingir todas as sinapses. Assim, a informação contextual é tratada naturalmente pela rede neural, considerando não somente cada padrão de entrada, mas também a correlação entre os padrões.
- **Tolerância a Falhas:** Se um neurônio um conjunto de conexões apresentam ruídos, a recuperação de um padrão armazenado é prejudicada em qualidade. Devido à natureza distribuída das redes neurais, elas tendem a não apresentar falhas catastróficas quando perdem alguns de seus neurônios, apresentando somente uma degradação em seu desempenho.
- **Robustez:** As redes neurais processam ruídos ou informações incompletas de forma eficiente, o que lhes permite manter o desempenho mesmo quando algumas de suas conexões são desativadas.

- **Implementação em VLSI:** O poder de processamento paralelo de uma rede neural permite que elas resolvam rapidamente certas tarefas. Além disso, essa mesma característica faz com que a redes neurais sejam ótimas candidatas para implementação em circuitos integrados do tipo VLSI (*Very Large Scale Integration*), que possui a principal vantagem de fornecer um meio de se obter comportamentos complexos de uma forma hierárquica (Mead, 1989).
- **Uniformidade de Análise e Projeto:** No contexto das redes neurais, pode-se dizer que elas desfrutam de universalidade como processadores de informação, uma vez que a mesma notação é adotada em todos os domínios que envolvem a aplicação das mesmas. Esta uniformidade pode ser notada de diferentes maneiras. Basicamente, um neurônio é a unidade comum a todas as redes neurais. Outro ponto importante desta característica é a possibilidade de compartilhamento de teorias e algoritmos de aprendizagem por diferentes aplicações de redes neurais. E há ainda a possibilidade de se desenvolver redes modulares, através de uma integração homogênea de módulos, os já citados e*nsembles*.

3.6 Aplicações

A capacidade de aprendizagem e generalização das redes neurais artificiais permite que elas sejam aplicadas para a solução de diversos problemas da vida real, tanto daqueles ligados à pesquisa científico-tecnológica quanto aos negócios. O estudo de redes neurais é um campo multidisciplinar, já que a sua aplicação na solução dos diversos problemas envolvem a aquisição de informações de diversas áreas do conhecimento. Para Kohonen (1989), a contribuição das redes neurais artificiais se dá na área de representação do conhecimento e reconhecimento de padrões.

Atualmente, várias arquiteturas de redes neurais tem sido propostas de forma que as mesmas se adaptem a aplicações específicas. Segundo estudos apresentados durante o *IEEE Inter-national Conference On Neural Networks*, realizado

em 1994 em Washington, Estados Unidos, segue algumas das áreas de concentração em que as redes neurais são amplamente empregadas (Moreno, 2010):

- Análise de movimento (Mitra, 1994);
- Aprendizado e reconhecimento de padrões (Bartfai, 1995);
- Ciência cognitiva (Mahoney, 1994);
- Controladores inteligentes (Weierich and Von Rosenberg, 1994);
- Economia e negócios (Vishwakarma, 1994);
- Estimação e identificação (Kane and Milgran, 1994);
- Inteligência computacional (Takahashi et al, 1994);
- Neurobiologia (Heileman, Georgiopoulos and Hwang, 1994);
- Previsão de séries temporais (Englisch and Hiemstra, 1994);
- Processamento e transmissão de sinais (Specht and Romsdahl, 1994);
- Sistemas de energia (Tsai, Tai and Reynolds, 1994);
- Sistemas híbridos (Park, 1994).

4
Teoria da Ressonância Adaptativa (ART)

Neste capítulo são apresentados os principais conceitos sobre a Teoria da Ressonância Adaptativa. Primeiramente é apresentada uma visão histórica dos fatos que marcaram o desenvolvimento desta teoria. Em seguida, são apresentadas as principais características dos modelos ART e ARTMAP e alguns modelos neurais desta família de redes.

4.1 Histórico

A ressonância adaptativa é um fenômeno bastante observado na natureza e é um dos principais métodos empregados pelo cérebro humano no processo de reconhecimento. O conceito de ressonância corresponde a "oscilações" que ocorrem em um meio físico quando este é submetido a uma frequência específica, denominada frequência de ressonância. Nas redes neurais baseadas na Teoria da Ressonância Adaptativa, a aprendizagem se dá quando as informações, na forma de saída dos neurônios, "oscilam" entre as camadas da rede desenvolvendo um estado de ativação de equilíbrio, equivalente neural à ressonância em um sistema físico (Amorim, 2006).

Sua origem pode ser atribuída a Von der Malsburg, que em 1973 apresentou uma rede neural artificial com características auto-organizáveis baseada no córtex visual dos maiores verte-brados. O trabalho de Malsburg influenciou o pensamento de Teuvo Kohonen, que desenvolveu os mapas auto-organizáveis em 1982 (Kohonen, 1982).

Influenciou também o trabalho de Steven Grossberg (Grossberg, 1976) que buscava uma solução para o dilema estabilidade-plasticidade. A plasticidade é a capacidade de um sistema aprender coisas novas, enquanto que a estabilidade

garante que o conhecimento adquirido seja preservado, de forma que este conhecimento não seja superposto pelos novos padrões. Grossberg resolveu tal problema com a proposta da primeira rede neural artificial baseada na teoria da ressonância adaptativa, denominada GN (*Grossberg Network*). A rede GN apresenta duas camadas. A primeira camada é responsável pela normalização dos dados de entrada. Já a segunda camada é uma camada competitiva com realimentação não-linear, responsável por selecionar os padrões de saída e os armazenar em uma memória de curto prazo.

A parceria de Gail Carpenter e Grossberg, em 1987, resultou no modelo neural ART1 (Carpenter & Grossberg, 1987a). Esta rede é caracterizada pelo treinamento não-supervisionado e possui a capacidade de agrupar padrões binários. A arquitetura ART1 é o modelo mais conhecido da família ART, tanto em termos de pesquisa relacionadas quanto em aplicações. No mesmo ano, Carpenter e Grossberg também propuseram o modelo ART2 (Carpen-ter & Grossberg, 1987b). Neste modelo, que também possui treinamento não-supervisionado, estende a arquitetura ART1 para trabalhar também com padrões analógicos.

A rede ARTMAP foi proposta em 1991 e é composta por dois módulos ART1, denominados ART_a e ART_b, que são interligados por um módulo inter-ART responsável por "casar" as informações de entrada e saída codificadas nos módulos ART_a e ART_b (Carpenter et al., 1991a). Foi a primeira arquitetura baseada na teoria da ressonância adaptativa apropriada para o reconhecimento de padrões. Apresenta treinamento supervisionado e auto-organizável e foi fundamental para o desenvolvimento de todas as outras redes ART que apresentam treinamento supervisionado.

Ainda em 1991, foi proposta também a rede neural ART *Fuzzy* (Carpenter et al. 1991b). Essa rede possui treinamento não-supervisionado e seus cálculos baseados na lógica *Fuzzy* (ou nebulosa) permitem o agrupamento de padrões de entrada com valores reais. Em 1992, é proposta a ARTMAP *Fuzzy*. Esta extensão do modelo ARTMAP permitiu a classificação de padrões com valores reais através de uma arquitetura de rede supervisionada (Carpenter et al. 1992).

Desde então, diversas arquiteturas de redes ART foram e tem sido propostas, a fim de se otimizar o processo de generalização destes modelos.

4.2 Características Básicas do Sistema ART

As redes neurais artificiais baseadas na Teoria da Ressonância Adaptativa são sistemas competitivos não-supervisionados que tem a capacidade de classificar padrões de entrada em categorias de conhecimento. Em um modelo ART, a regra de similaridade é determinada pelo grau de semelhança entre um padrão previamente armazenado e define a qual classe um padrão deve pertencer. Uma característica importante das redes neurais da família ART consiste em que se nenhuma das classes ressoa com um dado padrão de entrada, então uma nova classe é criada de acordo com um parâmetro de vigilância. Este parâmetro de vigilância é escolhido previamente e pode ser calibrado, o que permite ao usuário controlar o grau de similaridade dos padrões que pertencem a uma mesma classe (Carpenter & Grossberg, 1987).

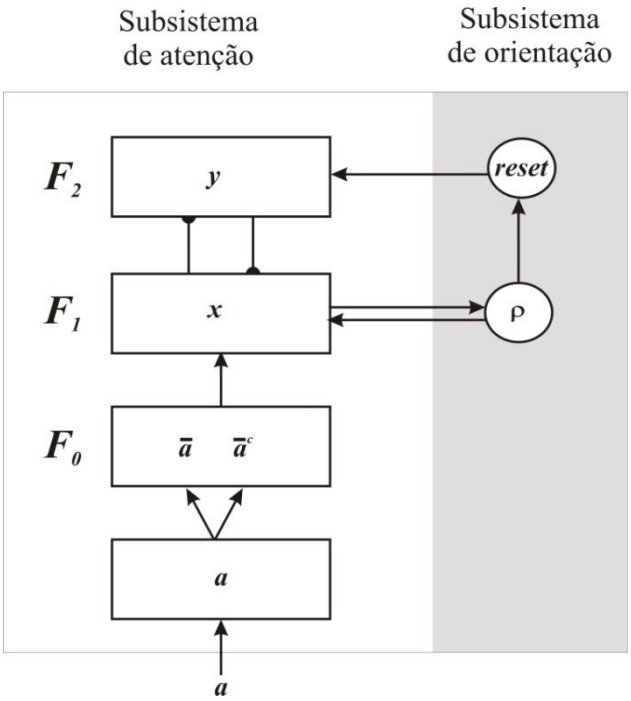


Figura 7 - Arquitetura básica do sistema ART.

A arquitetura básica de uma rede neural ART é formada por dois subsistemas: o subsistema de atenção e o subsistema de orientação, que podem ser observados na Figura 7. O subsistema de atenção é formado por duas camadas de neurônios interconectados. A camada F_1 é responsável pelo processamento dos

dados de entrada, ao passo que a camada F_2 agrupa os padrões em categorias de reconhecimento. No subsistema de orientação, são realizados a ressonância e o *reset*, utilizando um parâmetro de vigilância ρ , determina se um padrão de entrada pode ser incluído em uma das classes já existentes.

A camada de entrada F_0 é constituída por neurônios que pré-processam os vetores de entrada *a*. Este pré-processamento é denominado etapa de codificação e consiste na combinação das operações de normalização e complementação dos dados, a fim de que todos os padrões de entrada passem a ter a mesma magnitude. Após a etapa de codificação, os valores da camada F_0 são enviados à camada F_1 onde será realizada a segunda fase do treinamento ART, conhecida como etapa de reconhecimento.

Em busca de se controlar a similaridade dos padrões de entradas associados a uma mesma categoria, os modelos ART apresentam dois grupos de conexões entre as camadas F_1 e F_2 .Nas conexões não-recorrentes ou diretas (*feedforward*) de F_1 a F_2 , cada neurônio de F_1 se conecta a todos os neurônios de F_2 com pesos associados de baixo para cima. Já nas conexões recorrentes (*feedback*), observa-se a conexão de cada neurônio de F_2 a todos os neurônio de F_1, em conexões de cima para baixo, logo após a ressonância dos pesos.

É interessante notar que a memória de curto prazo dos sistemas ART é desenvolvido sobre os neurônios das camadas F_1 e F_2. Já a memória de longo prazo do sistema é associado às sinapses ascendentes e descendentes estabelecidas entre estas camadas (Lopes, 2005).

Na camada F_2, sempre que o neurônio representativo de cada categoria retorne a máxima entrada ponderada a partir das conexões ascendentes, este se torna o candidato ao armazenamento do conhecimento relativo ao dado padrão de entrada. Em consequência, todos os outros neurônios de F_2 são anulados. No próximo passo, os neurônios de F_1 combinam informações adquiridas nos processamentos efetuados em F_0 e F_2. Este fato é denominado etapa de comparação e representa a terceira fase de treinamento em uma rede neural ART.

A quarta fase do treinamento ART, também chamada de etapa de busca, é quando a efetivação do aprendizado é realizada. É neste momento que acontece a incorporação do padrão de entrada ao neurônio candidato escolhido em F_2, caso o neurônio candidato e o padrão apresentado sejam suficientemente similares. O

reset é responsável por tomar a decisão, e leva em consideração os sinais recebidos de F_1 e F_2. Esta fase do treinamento é um processo iterativo, e acontece até que um neurônio seja escolhido como vencedor. De outra maneira, se um neurônio candidato não for representativo para o padrão de entrada ele é inibido e busca-se um novo neurônio de F_2, que passa a ser o novo candidato, até que um neurônio seja o vencedor.

De maneira bem sucinta, pode-se dizer que o subsistema de atenção é responsável pela indicação de uma classe vencedora. Já o subsistema de orientação tem a responsabilidade de aceitar a classe proposta ou dizer que uma nova categoria deve ser buscada (Granger et al., 1997).

4.2.1 Parâmetros de uma Rede ART

Alguns parâmetros são de extrema importância para o desenvolvimento do aprendizado em modelos ART. A seguir, são apresentados estes parâmetros e como atuam na dinâmica do modelo:

- **Taxa de aprendizagem** α**,** $\alpha > 0$**:** o valor deste parâmetro indica a representatividade que os novos padrões pertencentes a cada categoria terão na atualização dos pesos no próximo ciclo do treinamento.
- **Taxa de treinamento** β**,** $0 \leq \beta \leq 1$**:** um valor de β próximo a zero indica um treinamento lento, ao passo que, com um valor próximo a um tem-se um treinamento rápido.
- **Parâmetro de vigilância** ρ**,** $0 \leq \rho \leq 1$**:** este parâmetro verifica a combinação dos padrões de entrada e dos pesos para que ocorra a ressonância, além de indicar quantas classes serão criadas. O número de classes varia de acordo com o valor adotado para ρ.

Ressalta-se que as redes neurais ART são muito sensíveis a variações em seus parâmetros, principalmente em relação ao parâmetro de vigilância (ρ) (Silva e Moreno, 2012).

Segundo Moreno (2010), o processo de atualização dos pesos na rede podem ser feitos por duas abordagens operacionais distintas. No treinamento sincronizado, a rede efetua um passo de propagação a cada inicialização e é

indicado quando se necessita estabelecer um caminho do padrão através da rede. No treinamento estável, a rede propaga o padrão até que se obtenha a classificação para o mesmo, o que faz com que a rede exija menos tempo para execução da tarefa.

4.2.2 Alguns modelos neurais ART:

- **Rede ART1:** Esta rede é baseada na Grossberg Network (GN) (Grossberg, 1976). Apresenta aprendizagem não-supervisionado de padrões de entrada binários. Em teoria, a rede ART1 (Carpenter & Grossberg, 1987a) é caracterizada por um conjunto de equações diferenciais não-lineares que são responsáveis por implementar as propriedades de estabilidade e plasticidade. Mas para efeito de implementação computacional o modelo pode ser resumido em um algoritmo sequencial comum.
- **ART *Fuzzy***: A rede neural ART *Fuzzy*, ou ART nebulosa (Carpenter & Grossberg, 1992) constitui uma evolução do modelo ART1. Este modelo utiliza cálculos da Teoria dos conjuntos nebulosos (Zadeh, 1965). A estrutura da rede ART *Fuzzy* é essencialmente a mesma, mas substitui o operador de interseção $\cap$ binário do modelo ART1 pelo operador AND nebuloso, representado por $\wedge$. Dessa maneira, o modelo ART1 que antes era restrito a trabalhar apenas com dados binários, passa a trabalhar também com dados analógicos. Esta alteração, junto com a codificação em complemento, preserva amplitude da informação e normaliza os vetores de entrada, além de garantir um modelo de classificação não-supervisionado com grande velocidade de aprendizagem (Amorim, 2006).
- **ART Euclidiana:** A rede neural ART Euclidiana (Kenaya, 2008) é um modelo neural que utiliza a distância Euclidiana entre os padrões de entrada e os respectivos centros de classes para escolher a qual classe o padrão pertencerá. A rede ART Euclidiana apresenta algumas diferenças em relação à ART *Fuzzy*. A primeira delas consiste na substituição do operador AND *Fuzzy* pela distância Euclidiana. Assim,

passa-se a ter hiper-esferas centradas no centro das classes ao invés dos hiper-retângulos criados pela lógica *Fuzzy*. Outra particularidade deste modelo é que ele não requer nem a normalização nem a codificação dos vetores de entrada, reduzindo o custo computacional. Mais uma característica particular deste modelo está na escolha das categorias. Enquanto na rede ART *Fuzzy* a categoria escolhida é aquela com ativação máxima, isto é, uma alta vigilância leva a uma rigorosa condição de ressonância e um maior número de nós de saída, na rede ART Euclidiana essas condições são alcançadas com a escolha da categoria com ativação mínima.

- **Redes ART *Fuzzy* Modificada** e **ART Euclidiana Modificada:** Tanto na rede ART *Fuzzy* quanto na rede ART Euclidiana tradicionais, o treinamento é iniciado com o número de neurônios igual ao número de padrões apresentados a rede. Assim sempre que estas redes vão procurar um neurônio para armazenar o conhecimento de um padrão de entrada, calcula-se a função de ativação para todos os neurônios existentes, tanto os ativos quanto os inativos. Nas redes ART *Fuzzy* Modificada e ART Euclidiana Modificada (Moreno, 2010), o treinamento é iniciado com apenas um neurônio ativo, de maneira que durante o treinamento o modelo se expande, criando novos neurônios para abrigar novos padrões. Este comportamento é semelhante ao que acontece no cérebro, onde há uma grande quantidade de neurônios inativos e que serão ativados sempre que os neurônios já ativos não sejam capazes de armazenar o novo conhecimento.

- **Redes ART *Fuzzy* Modificada com Treinamento Continuado** e **ART Euclidiana Modificada com Treinamento Continuado:** As redes ART *Fuzzy* Modificada com Treinamento Continuado e ART Euclidiana Modificada com Treinamento Continuado apresenta dois diferenciais em relação às redes ART tradicionais (Moreno, 2010). A modificação na rede ART tradicional torna o processo de busca pelo neurônio vencedor mais rápido e isso sem que a rede perca sua eficiência. Com esta modificação, a rede é iniciada com apenas um cluster e durante o treinamento se expande para abrigar novos padrões. O módulo de

treinamento continuado permite que a rede neural consiga armazenar novos dados sem necessidade de um novo treinamento

4.3 Características Básicas do Sistema ARTMAP

Os modelos ART foram propostos originalmente para a classificação não-supervisionada de padrões. Entretanto, as redes neurais ARTMAP (Carpenter et al., 1991a) foram projetados a fim de se obter o treinamento supervisionado e auto-organizável sem perder a estabilidade do processo de treinamento. A arquitetura básica das redes ARTMAP pode ser observada na Figura 8.

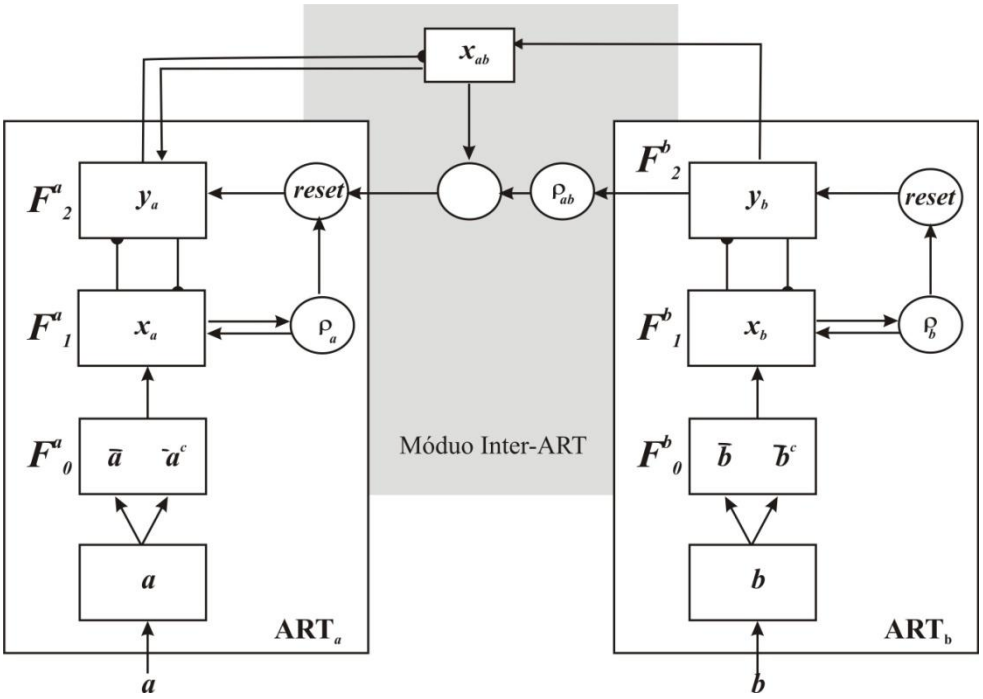


Figura 8 – Arquitetura básica do sistema ARTMAP.

A arquitetura ARTMAP é constituída por dois módulos ART, o módulo ART_a e o módulo ART_b , e um módulo *inter*-ART. Tanto o módulo ART_a como o módulo ART_b desempenham as funções de um módulo ART, como descrito anteriormente. Durante o treinamento, o módulo ART_a recebe um conjunto de padrões de entrada *a* e o módulo ART_b um conjunto de entradas *b*, sendo b o valor correto esperado associado a entrada a. Já o módulo *inter*-ART é responsável por

associar os clusters de entrada criados pelo módulo ART_a com os clusters de saída criados pelo módulo ART_b, de forma que um "casamento" seja estabelecido entre as entradas e saídas, ou seja, a vigilância do sistema.

É importante ressaltar que durante o treinamento de uma rede neural ARTMAP, tanto o módulo ART_a quanto o módulo ART_b trabalham *on-line*. Já na fase de análise, o módulo ART_b é desativado, e a análise é feita apenas com o módulo ART_a.

4.3.1 Parâmetros de uma Rede ARTMAP

Os parâmetros que devem ser configurados para o desenvolvimento do aprendizado dos modelos ARTMAP são os mesmos dos modelos ART. E já que os modelos ARTMAP são compostos por dois módulos ART e um módulo *inter*-ART, é necessária a definição de um parâmetro de vigilância para cada um desses módulos.

- Taxa de aprendizagem α, $\alpha > 0$;
- Taxa de treinamento β, $0 \leq \beta \leq 1$;
- Parâmetro de vigilância do módulo ART_a ρ_a, $0 \leq \rho_a \leq 1$;
- Parâmetro de vigilância do módulo ART_b ρ_b, $0 \leq \rho_b \leq 1$;
- Parâmetro de vigilância do módulo inter-ART ρ_{ab}, $0 \leq \rho_{ab} \leq 1$.

4.3.2 Alguns modelos ARTMAP:

- **Rede ARTMAP1:** É composta por dois módulos ART1 e, portanto, utiliza o operador de interseção binário para o cálculo da função de ativação e do teste de vigilância. Como é um modelo supervisionado composto por dois módulos ART1, sua aplicação se restringe a problemas em que os padrões de entrada são de natureza binária.
- **ARTMAP *Fuzzy*:** A rede neural ARTMAP *Fuzzy* (Carpenter & Grossberg, 1992) é uma evolução do modelo ARTMAP1, composta por

dois módulos ART *Fuzzy*. De maneira análoga ao modelo ART *Fuzzy*, cada um dos módulos ART *Fuzzy* realiza seus cálculos na lógica nebulosa, permitindo que este modelo aceite padrões de entrada analógicos.

- **ARTMAP Euclidiana:** Composta por dois módulos ART Euclidiana, sendo que cada um desses módulos preservam as características do modelo não-supervisionado.
- **Redes ARTMAP *Fuzzy* Modificada** e **ARTMAP Euclidiana Modificada:** Estes modelos são compostos por dois módulos ART *Fuzzy* Modificada e ART Euclidiana Modificada, respectivamente. As diferenças destes modelos em relação à ARTMAP *Fuzzy* e ARTMAP Euclidiana são as mesmas observadas nos modelos não-supervisionados, o que garante maior eficiência no treinamento.

5
Software e Resultados

Neste capítulo são apresentados o software e os resultados obtidos por ele no reconhecimento de padrões de algumas bases de dados. Na seção 5.1 é feita a descrição do aplicativo, bem como do padrão de arquivo a ser usado como entrada para o mesmo. Na seção 5.2 são discutidos os resultados obtidos para quatro base de dados, a saber, Wisconsin Diagnostic Breast Cancer, Vertebral Column Data Set, Oxford Parkinson's Disease Detection Dataset e Movement Libras.

5.1 Software

Este trabalho apresenta um software simples e de fácil utilização para se realizar reconhecimento de padrões. As redes neurais foram implementadas na linguagem de programação JAVA e a estrutura do software pode ser observado na Figura 9. Os componentes do software são:

- Base de dados;
- Núcleo;
- Interface.

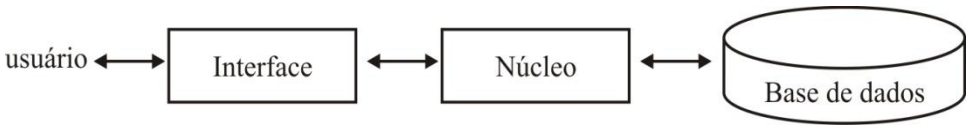


Figura 9 - Estrutura do Software proposto.

Os componentes serão detalhados nas próximas seções.

5.1.1 Base de dados

Os dados de entrada a serem apresentados aos modelos são adquiridos a partir de arquivos texto, tanto para o treinamento quanto para a análise. O formato do arquivo de entrada de dados do software pode ser visualizado na Figura 10.

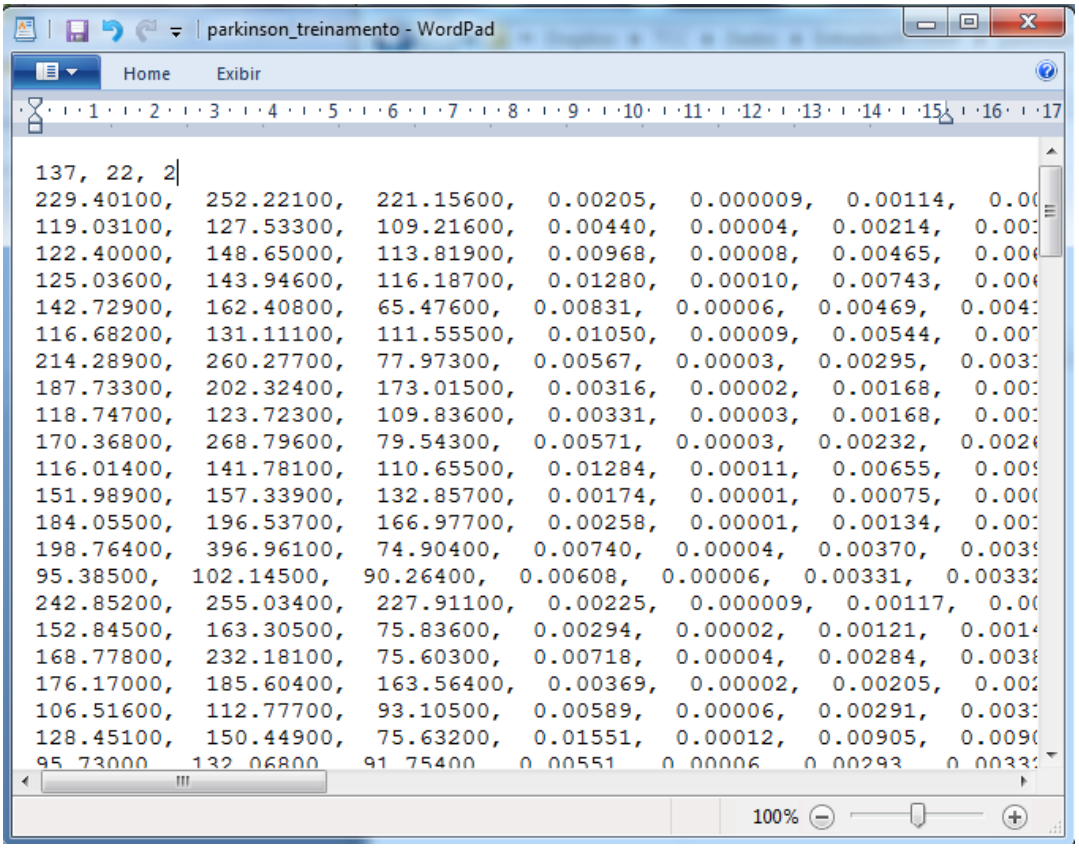


Figura 10 – Exemplo de Arquivo de entrada.

Os arquivos texto devem seguir o seguinte padrão:

- Não existe nenhum cabeçalho no arquivo;
- Na primeira linha estão o número de padrões a serem apresentados ao modelo, o número de características que compõe cada padrão e o número de classes em que estes padrões devem ser classificados;
- Cada uma das linhas a partir da segunda representa um padrão a ser classificado;
- Os valores reais devem ser separados por ponto;
- Cada valor deve ser separado por vírgula.

5.1.2 Núcleo

Neste componente estão implementados os modelos neurais escolhidos. Aqui são obtidos os padrões de entrada para os modelos, tanto do conjunto de treinamento quanto do conjunto de análise, e são realizadas as fases de treinamento e análise destas entradas. Ou seja, o software recebe algumas características do objeto a ser reconhecido através do conjunto de treinamento e realiza o mesmo, a fim de se desenvolver o conhecimento. Logo após é realizada a análise, que retornará a classificação para cada um dos padrões do conjunto de testes.

5.1.3 Interface

A fim de permitir que qualquer usuário utilize o aplicativo, independentemente do grau de conhecimento sobre redes neurais artificiais, a interface para execução de mesma é bem simples, e pode ser visualizada na Figura 11.

Primeiramente o usuário deve escolher qual dos modelos neurais deve-se executar. Em seguida, deve-se configurar os parâmetros, que variam de acordo com os modelo escolhido. O terceiro passo consiste na escolha dos arquivos de dados que serão utilizados como treinamento e teste. Isto pode ser feito clicando nos botões "Entrada Treinamento" e "Entrada Teste", e desta forma uma nova janela será aberta, permitindo ao usuário navegar em seus diretórios e buscar pelos respectivos arquivos. Estes campos também variam em função da rede escolhida, já que as redes ART *Fuzzy* Modificada com Treinamento Continuado e ART Euclidiana Modificada com Treinamento Continuado utilizam apenas um arquivo de entrada, composto pelos padrões de treinamento e de teste.

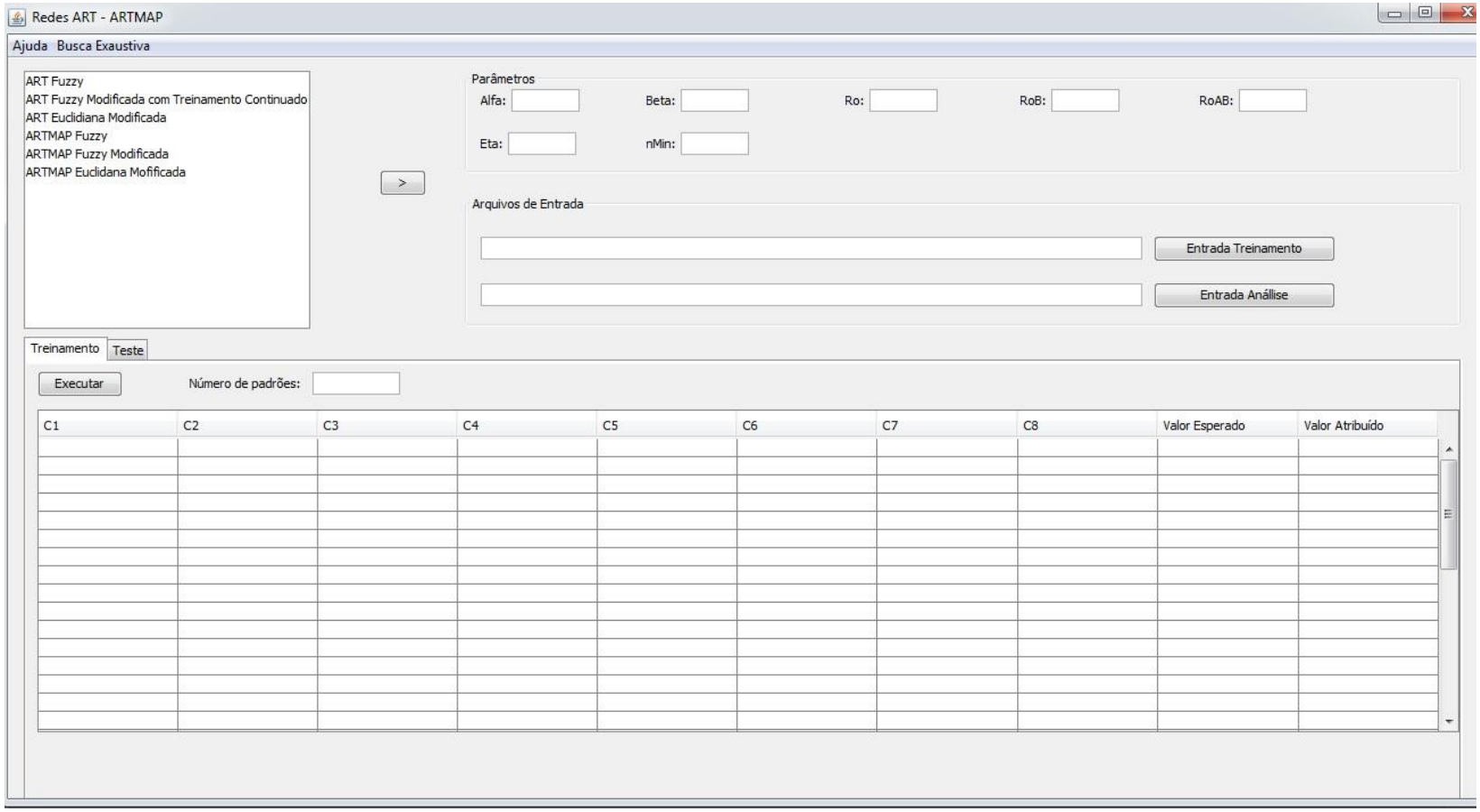


Figura 11 – Tela principal do aplicativo.

A parte inferior da tela apresentada na Figura 11 possui duas abas, intituladas "Treinamento" e "Teste". Nestas duas abas são apresentadas tabelas com os dados carregados a partir dos arquivos de entrada, sendo que cada linha representa um padrão a ser apresentado a rede, e cada coluna representa uma característica do padrão. Além das características, a coluna "Classe Esperada" e "Classe Atribuída" representam a real classe do padrão, atribuída pela natureza do problema em avaliação, e a classe atribuída durante o processo de classificação, respectivamente. Tanto o treinamento da rede neural quanto a execução dos testes podem ser executados através dos botões "Executar" das respectivas abas. A aba testes, além das características já apresentadas, apresenta os campos "Número de Acertos" e "% de Acertos", que informam ao usuário o quanto o modelo escolhido acertou no processo de classificação.

A outra funcionalidade do aplicativo é permitir ao usuário realizar buscas exaustivas, de modo que o mesmo possa avaliar o melhor conjunto de parâmetros para uma dada base de dados. Na Figura 12 está representada a tela que permite o usuário configurar esta busca, que pode apenas ser exibida em tela ou ser salva em um arquivo texto para posterior análise. Como já comentado no capítulo sobre as arquiteturas neurais baseados na Teoria da Ressonância Adaptativa, os parâmetros devem estar compreendidos nos seguintes intervalos:

- α, $\alpha > 0$;
- $0 \leq \beta \leq 1$;
- $0 \leq \rho \leq 1$.

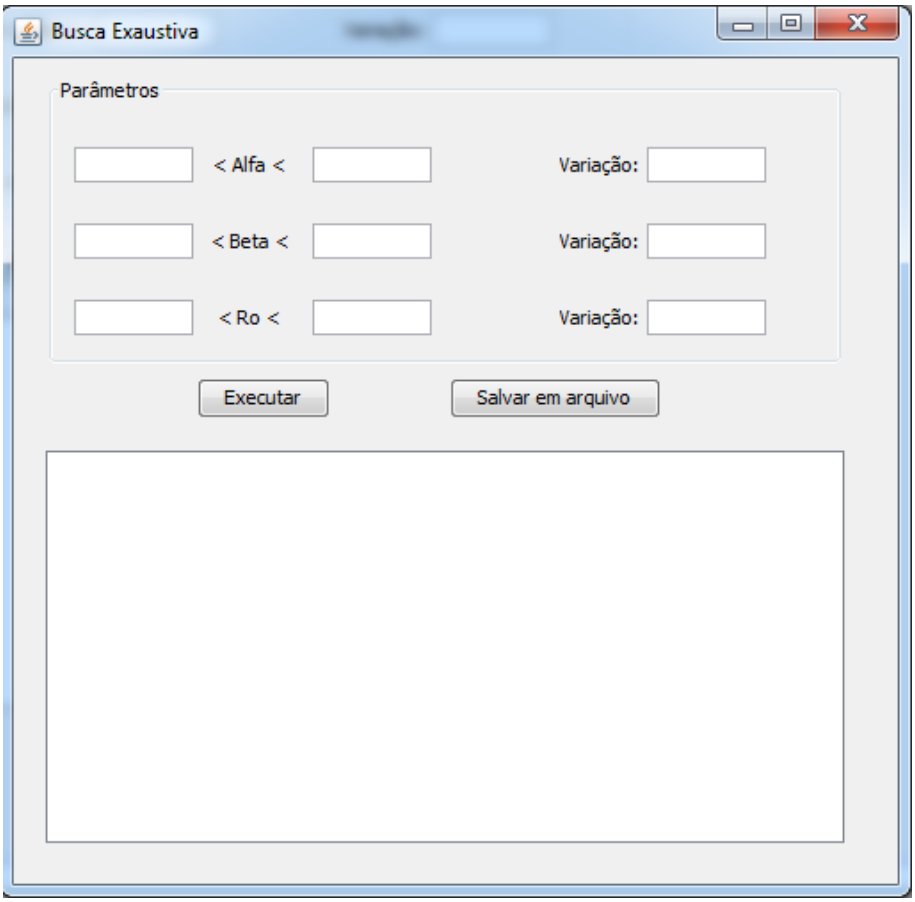


Figura 12 – Configuração da busca exaustiva.

5.2 Experimentos

Nesta seção são apresentados os resultados experimentais obtidos com o aplicativo. Foram feitos experimentos com quatro bases de dados disponíveis no repositório UCI - *Machine Learning Repository* disponível em <http:archive.ics.uci.edu/ml/>. Cada uma das bases de dados foi embaralhada utilizando um software e em seguida cada base foi dividida em dois conjuntos: o conjunto de treinamento com 70% dos dados e o conjunto de análise, com 30% dos dados. Os mesmos arquivos de entrada foram utilizados na execução de cada uma das redes neurais que compõe o software desenvolvido. Com o objetivo de se encontrar as melhores combinações para os parâmetros, foram realizadas buscas exaustivas. Para a discussão dos resultados são apresentadas as matrizes de confusão gerada com os

resultados obtidos pelo software. Uma matriz de confusão é uma maneira de estimar o número correto e incorreto de predições feito pelo classificador. Desta maneira, na diagonal principal da matriz estão representados as classificações corretas e os valores que não se encontram na diagonal principal representam as classificações incorretas (Beckmann, 2010).

5.2.1 Wisconsin Diagnostic Breast Cancer

Este banco de dados de câncer de mama foi desenvolvido a partir de dados do Hospital da Universidade de Wisconsin. Os dados foram coletados por Mangasariam e Wolberg (1990), que utilizaram imagens digitalizadas com a técnica "*fine needle aspirate*" para seios. São 569 padrões descritos por 30 atributos reais, que descrevem as características do núcleo da célula, calculados com base nas imagens digitalizadas. Destes 115 foram utilizados como conjunto de testes na análise da rede. A seguir são apresentadas as matrizes de confusão dos resultados obtidos.

Tabela 1 - Matriz de confusão da análise obtida pela rede ART *Fuzzy* referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	24	14
Benigno	13	64

A Tabela 1 apresenta a previsão feita pela rede ART *Fuzzy*, com parâmetros α = 0.1, β = 0.3 e ρ =0.993. Com esta rede, obteve-se um total de 88 acertos, ou seja, 76.52%.

Tabela 2 - Matriz de confusão da análise obtida pela rede ART *Fuzzy* Modificada com Treinamento Continuado referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	37	3
Benigno	0	75

Utilizando a rede ART *Fuzzy* Modificada com Treinamento Continuado com parâmetros α = 0.1, β = 0.1 e ρ =0.999, η =0.9 e *nMin* = 5 obteve-se um total de 112

acertos, o que representa 97.39% de classificações corretas. Na Tabela 2 está representada a matriz de confusão obtida com a classificação com esta rede.

Tabela 3 - Matriz de confusão da análise obtida pela rede ART Euclidiana Modificada referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	29	13
Benigno	8	65

Com a rede ART Euclidiana Modificada foram obtidos 94 acertos, isto é, 81.74% de acertos na classificação. A Tabela 3 apresenta a matriz de confusão para este resultado, com os parâmetros β = 0.7 e ρ =0.0001399.

Tabela 4 - Matriz de confusão da análise obtida pela rede ARTMAP *Fuzzy* referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	28	23
Benigno	9	55

Na Tabela 4 é apresentada a matriz de confusão para a previsão através da rede ARTMAP *Fuzzy*, com parâmetros α = 0.1, β = 0.1, ρ_a =0.85, ρ_b = 1.0 e ρ_{ab} 0.99. Neste experimento obteve-se um total de 83 acertos, o que representa 72.17% de previsões corretas.

Tabela 5 - Matriz de confusão da análise obtida pela rede ARTMAP *Fuzzy* Modificada referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	11	6
Benigno	26	72

A matriz de confusão apresentada na Tabela 5 é o resultado obtido pela classificação utilizando a rede ARTMAP *Fuzzy* Modificada, com parâmetros α = 0.1, β = 0.999, ρ_a =0.9, ρ_b = 1.0 e ρ_{ab} = 0.99. Esta rede acertou 83 dos 115 padrões utilizados para o teste, totalizando 72.17% de previsões corretas.

Tabela 6 - Matriz de confusão da análise obtida pela rede ARTMAP Euclidiana Modificada referente à base de dados *Wisconsin Diagnostic Breast Cancer.*

Esperado/Previsão	*Maligno*	*Benigno*
Maligno	11	1
Benigno	26	77

Utilizando a rede ARTMAP Euclidiana Modificada com parâmetros β = 0.2, ρ_a =0.25, ρ_b = 0.001 e ρ_{ab} = 0.01, obteve-se um total de 88 acertos, ou seja, 76.52% de classificações corretas. A matriz de confusão para este experimento é apresentada na Tabela 6.

Finalmente, na Tabela 7 apresenta-se de forma resumida os resultados obtidos no reconhecimento do diagnóstico das células de câncer da base de dados *Wisconsin Diagnostic Breast Cancer* obtidos neste trabalho, bem como alguns resultados disponíveis na literatura.

Tabela 7 – Resultados obtidos para a base de dados *Wisconsin Diagnostic Breast Cancer.*

Algoritmo	*Taxa de acertos (%)*	*Referência*
KDD	87.30	Shadvar, (2012)
SVM	86.9	Shadvar, (2012)
MLP	97.59	Freitas et. al., (2012)
ART *Fuzzy*	76.52	
ART *Fuzzy* MTC	97.39	
ART Euclidiana Modif.	81.74	
ARTMAP *Fuzzy*	72.17	
ARTMAP *Fuzzy* Modif.	72.17	
ARTMAP Euclidiana Modif.	76.52	

Como pode-se observar na tabela anterior e no gráfico da Figura 13, o melhor resultado obtido pelo software foi com a rede neural ART *Fuzzy* Modificada com Treinamento Continuado, que obteve 97.39% de acerto, um resultado muito próximo da literatura que é de 97.59%. Entretanto nesta rede o treinamento foi *on-line* enquanto que o resultado na literatura foi obtido *off-line.*

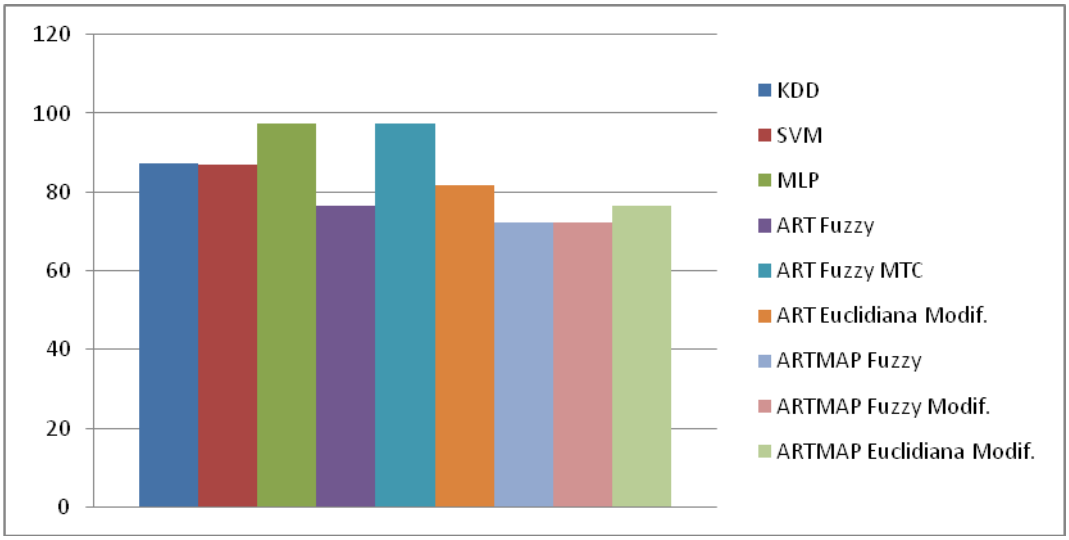


Figura 13 – Gráfico comparativo para os resultados obtidos com a base de dados *Wisconsin Diagnostic Breast Cancer.*

5.2.2 Vertebral Column Data Set

Base de dados construída por um brasileiro, Dr. Henrique da Mota, no Centro Médico Cirúrgico de Readaptação de Massues, na França. São 310 entradas que devem ser classificadas em três categorias: normal (N), hérnia de disco (H) e Espondilolistese (E). Cada padrão a ser classificado é representado por seis atributos derivados da forma e orientação da pelve e da coluna lombar. O conjunto de testes é constituído por 93 padrões e os resultados são discutidos a seguir.

Tabela 8 Matriz de confusão da análise obtida pela rede ART *Fuzzy* referente à base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	31	0	0
H	1	0	0
N	13	16	32

Na Tabela 8 está representado o resultado obtido pela rede ART *Fuzzy*, com parâmetros α = 0.1, β = 0.799 e ρ =0.97. Neste caso, obteve-se um total de 63 acertos, o que representa 67.74% de predições corretas.

Tabela 9 - Matriz de confusão da análise obtida pela rede ART *Fuzzy* Modificada com Treinamento Continuado referente à base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	44	1	0
H	1	15	0
N	0	0	32

Utilizando a rede ART *Fuzzy* Modificada com Treinamento Continuado obteve-se um total de 91 acertos, o que representa 97.85% de classificações corretas, com parâmetros α = 0.1, β = 0.1 e ρ =0.999, η =0.9 e *nMin* = 5. Na Tabela 9 está representada a matriz de confusão obtida com a classificação com esta rede.

Tabela 10 - Matriz de confusão da análise obtida pela rede ART Euclidiana Modificada referente *à* base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	34	3	9
H	1	10	3
N	10	3	20

Com a rede ART Euclidiana Modificada foram obtidos 68.82% de acertos, isto é, 64 dos 93 padrões utilizados no teste foram classificados corretamente. A Tabela 10 apresenta a matriz de confusão para este experimento, com os parâmetros β = 0.6 e ρ =0.00125.

Tabela 11 - Matriz de confusão da análise obtida pela rede ARTMAP *Fuzzy* referente à base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	45	1	6
H	0	15	26
N	0	0	0

Na Tabela 11 é apresentada a matriz de confusão para a previsão através da rede ARTMAP *Fuzzy,* com parâmetros α = 0.5, β = 0.89, ρ_a =0.8, ρ_b = 1.0 e

ρ_{ab} 0.99. Neste experimento obteve-se um total de 60 acertos, o que representa 64.52% de previsões corretas.

Tabela 12 - Matriz de confusão da análise obtida pela rede ARTMAP *Fuzzy* Modificada referente à base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	45	12	28
H	0	4	4
N	0	0	0

A Tabela 12 apresenta a matriz de confusão para a previsão através da rede ARTMAP *Fuzzy* Modificada, com parâmetros α = 0.799 β = 0.99, ρ_a =0.8, ρ_b = 1.0 e ρ_{ab} 0.99. O número de acertos com este modelo neural é de 49, ou seja, 52.69%.

Tabela 13 - Matriz de confusão da análise obtida pela rede ARTMAP Euclidiana Modificada referente à base de dados *Vertebral Column Data Set.*

Esperado/Previsão	*E*	*H*	*N*
E	44	16	32
H	1	0	0
N	0	0	0

Já com a rede ARTMAP Euclidiana Modificada com parâmetros β = 0.2, ρ_a = 0.24, ρ_b = 0.001 e ρ_{ab} = 0.01, obteve-se um total de 44 acertos, ou seja, 47.31% de classificações corretas. A matriz de confusão para este caso é apresentada na Tabela 13.

A comparação entre os resultados obtidos neste trabalho com alguns disponíveis na literatura pode ser observado na Tabela 14 e na Figura 14.

Tabela 14 - Resultados obtidos para a base de dados *Vertebral Column Data Set.*

Algoritmo	*Taxa de acertos (%)*	*Referência*
SVM	85.9	Neto et. al. (2011)
MLP	87.23	Freitas et. al. (2012)
ART *Fuzzy*	67.74	
ART *Fuzzy* MTC	97.85	
ART Euclidiana Modif.	68.82	
ARTMAP *Fuzzy*	64.52	
ARTMAP *Fuzzy* Modif.	52.69	
ARTMAP Euclidiana Modif.	47.31	

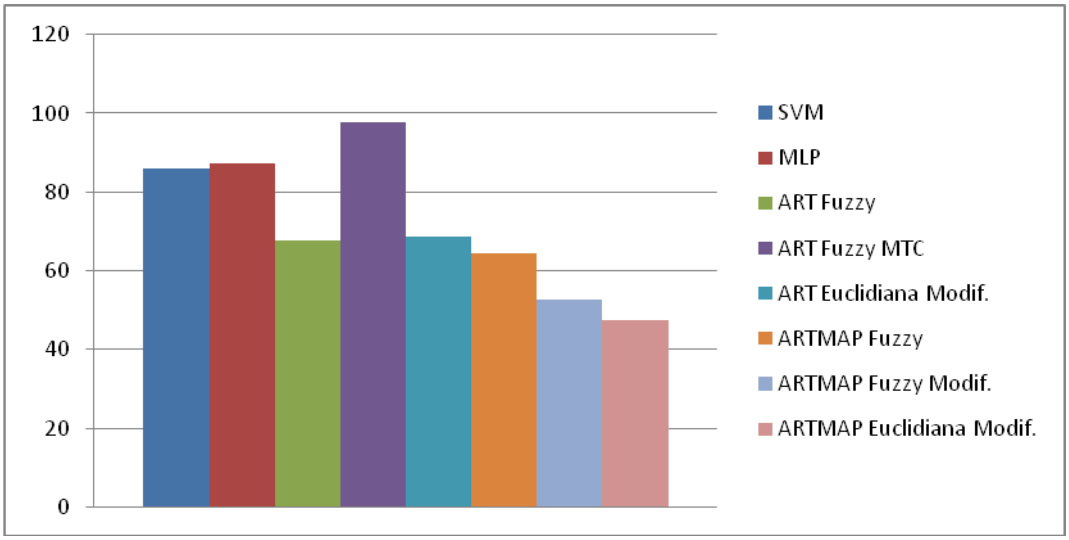


Figura 14 – Gráfico comparativo para os resultados obtidos para a base de dados *Vertebral Column Data Set.*

O resultado obtido pelo aplicativo foi melhor do que o encontrado na literatura, isto pode ser observado tanto na tabela quanto no gráfico acima.

5.2.3 Oxford Parkinson's Disease Detection Dataset

Este conjunto de dados foi criado na Universidade de Oxford, em colaboração com o Centro Nacional de Voz e Fala de Denver - Colorado, onde foram gravados os sinais de fala. O estudo original (Little et. al., 2007) apresenta os métodos de extração utilizados para extração das características da voz. O conjunto de dados é composto de uma série de medições biomédicas da voz de 31 pessoas, 23 com

doença de Parkinson. São cerca de seis gravações por paciente e o objetivo é discriminar pessoas saudáveis daquelas com doença de Parkinson (DP). São 197 padrões, caracterizados por 22 atributos. A seguir são apresentadas as respectivas matrizes de confusão dos experimentos realizados com as redes neurais implementadas neste trabalho.

Tabela 15 - Matriz de confusão da análise obtida pela rede ART *Fuzzy* referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	43	14
sem DP	0	1

Utilizando a rede ART *Fuzzy*, com parâmetros α = 0.1, β = 0.79 e ρ =0.989 obteve-se um total de 44 acertos, o que representa 75.86% de predições corretas. A matriz de confusão é apresentada na Tabela 15.

Tabela 16 - Matriz de confusão da análise obtida pela rede ART *Fuzzy* Modificada com Treinamento Continuado referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	41	2
sem DP	2	13

Com a rede ART *Fuzzy* Modificada com Treinamento Continuado obteve-se um total de 54 acertos, ou seja, 93.10% das classificações foram corretas. Os parâmetros utilizados foram α = 0.1, β = 0.899 e ρ =0.999, η =0.9 e *nMin* = 5 e a matriz de confusão pode ser observada na Tabela 16.

Tabela 17 – Matriz de confusão obtida pela rede ART Euclidiana Modificada referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	43	15
sem DP	0	0

A Tabela 17 apresenta a matriz de confusão obtida com a classificação feita com a rede ART Euclidiana Modificada com os parâmetros β = 0.1 e ρ =0.0009.

Observa-se que foram feitas 43 predições corretas, o que representa 74.14% de acertos.

Tabela 18 - Matriz de confusão obtida pela rede ARTMAP *Fuzzy* referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	43	15
sem DP	0	0

A Tabela 18 apresenta a matriz de confusão para a previsão através da rede ARTMAP *Fuzzy* para a base de dados *Oxford Parkinson's Disease Detection Dataset.* Os parâmetros utilizados foram α = 0.799 β = 0.99, ρ_a =0.8, ρ_b = 1.0 e ρ_{ab} 0.99 e o número de classificações corretas com este modelo neural foi de 43, ou seja, 74.14%.

Tabela 19 - Matriz de confusão obtida pela rede ARTMAP *Fuzzy* Modificada referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	43	14
sem DP	0	1

O número de predições corretas obtidas com a rede ARTMAP *Fuzzy* Modificada é de 44, com os parâmetros utilizados foram α = 0.6 β = 0.6, ρ_a =0.8, ρ_b = 1.0 e ρ_{ab} 0.99. Neste caso, a rede acertou em 75.86% e respectiva matriz de confusão pode ser observada na Tabela 19.

Tabela 20 - Matriz de confusão obtida pela rede ARTMAP Euclidiana Modificada referente à base de dados *Oxford Parkinson's Disease Detection Dataset.*

Esperado/Previsão	*com DP*	*sem DP*
com DP	2	1
sem DP	41	14

Para a base de dados testada nesta seção, a rede ARTMAP Euclidiana Modificada apresentou o pior resultado, acertando apenas 16 dos 58 padrões, isto é, 27.59% do conjunto de teste, isto pode ser observado na Tabela 20. Os parâmetros utilizados foram β = 0.2, ρ_a =0.25, ρ_b = 0.001 e ρ_{ab} 0.01.

Na Tabela 21 é apresentado um resumo dos resultados obtidos e alguns dos resultados obtidos na literatura para a base de dados *Oxford Parkinson's Disease Detection Dataset*.

Tabela 21 - Resultados obtidos para a base de dados *Oxford Parkinson's Disease Detection Dataset*.

Algoritmo	*Taxa de acertos (%)*	*Referência*
IS	79.78	Ene (2008)
MCS	80.92	Ene (2008)
HS	81.28	Ene (2008)
Regressão Logística	89.72	Anderson (2009)
MLP	81.03	
ART *Fuzzy*	75.86	
ART *Fuzzy* MTC	93.10	
ART Euclidiana Modif.	74.14	
ARTMAP *Fuzzy*	74.14	
ARTMAP *Fuzzy* Modif.	75.86	
ARTMAP Euclidiana Modif.	27.59	

Como pode ser observado na tabela anterior e no gráfico da Figura 15, os resultados obtidos pela rede ART *Fuzzy* Modificada com Treinamento Continuado obtiveram os melhores resultados, 93.10%, e, considerando que o aplicativo é executado em tempo real, o resultado sendo, inclusive, melhor do encontrado na literatura, 89.72%, cuja execução é *off-line*.

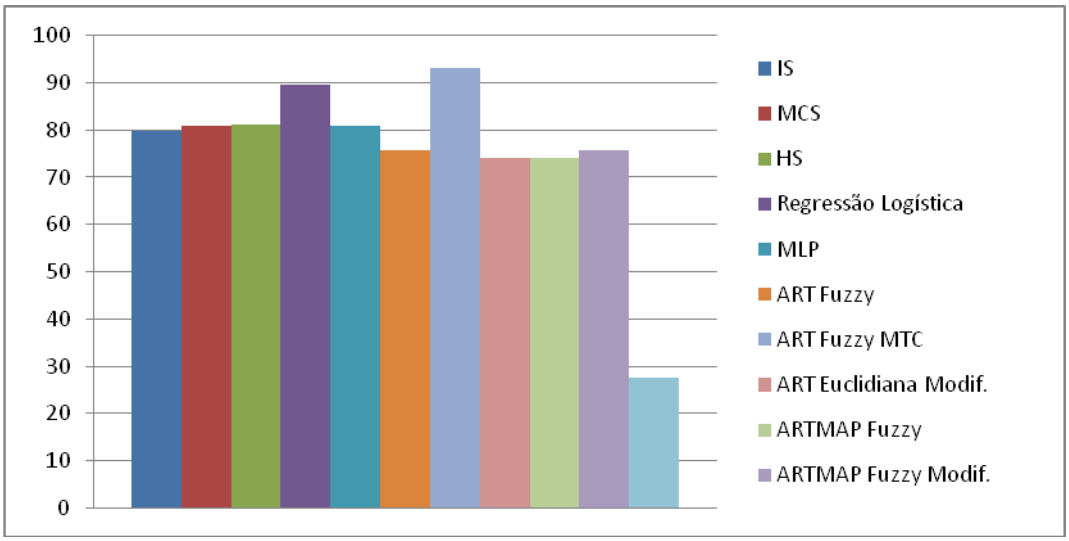


Figura 15 –Gráfico comparativo para os resultados obtidos para a base de dados *Oxford Parkinson's Disease Detection Dataset*.

5.2.4 Movement Libras

O conjunto de dados *Movement Libra*s contém 360 padrões que devem ser reconhecidos como um movimento da Língua Brasileira de Sinais. O movimento é representado como uma curva bidimensional que foi obtida a partir de vídeos. Cada vídeo corresponde a apenas um movimento da mão e tem cerca de 7 segundos. Cada movimento observado foi mapeado em uma representação com 90 características, e os 360 padrões devem ser reconhecidos como um dentre os 15 movimentos observados.

Como a matriz de confusão teria 15 linhas e 15 colunas (número de movimentos observados) optou-se por não apresentá-las. Os resultados obtidos para este experimento juntamente com os da literatura são observados na Tabela 22.

Tabela 22 – Resultados obtidos com a base de dados *Movement Libra*s.

Algoritmo	*Taxa de acertos (%)*	*Referência*
K-means	59.0	Amin & Sutariya (2013)
Inproved K-means	61.0	Amin & Sutariya (2013)
MLP	74.07	Freitas et. al. (2012)
RBF	75.92	Freitas et. al. (2012)
KNN	81.9	Shadvar (2012)
SVM	72.5	Shadvar (2012)
ART *Fuzzy*	12.03	
ART *Fuzzy* MTC	27.78	
ART Euclidiana Modif.	19.44	
ARTMAP *Fuzzy*	27.80	
ARTMAP *Fuzzy* Modif.	14.44	
ARTMAP Euclidiana Modif.	26.67	

Como pode ser observado na tabela acima e na Figura 16, as redes neurais não obtiveram bons resultados, isto se deve a alta dimensionalidade dos padrões de entrada (90 características). Como o objetivo deste trabalho era fazer só o reconhecimento dos padrões, não preocupou-se com a utilização de técnicas para a diminuição da dimensionalidade dos dados.

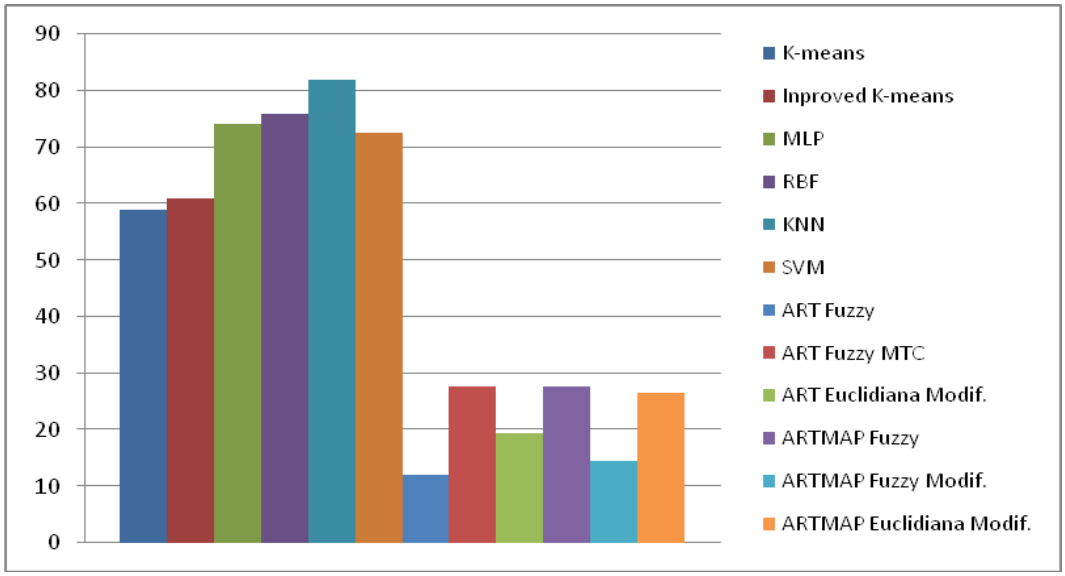


Figura 16 - Gráfico comparativo para os resultados obtidos com a base de dados *Movement Libras*.

6
Conclusões

Este capítulo apresenta as conclusões deste trabalho de conclusão de curso. É organizado da seguinte forma: Na Seção 6.1,são apresentas as conclusões gerais do trabalho, enquanto que na Seção 6.2, são apresentadas as sugestões de trabalhos futuros.

6.1 Conclusões

O software desenvolvido neste trabalho de conclusão de curso atingiu os objetivos iniciais do trabalho, isto é, desenvolver um software de fácil utilização e que qualquer usuário possa utilizar, sem necessidade de conhecer a natureza do processo de classificação. As redes implementadas no núcleo do software apresentaram resultados satisfatórios nos casos em que as bases de dados apresentam padrões com um número menor de atributos (até 30 atributos). No caso da base de dados em que os padrões tem maior dimensionalidade (90 atributos), o resultado não foi bom, mas como o objetivo do trabalho era se concentrar na classificação dos padrões, não se preocupou em realizar qualquer tratamento dos dados antes de submeter os mesmos as redes neurais artificiais.

6.2 Sugestões para trabalhos futuros

Os resultados obtidos neste trabalho de conclusão de curso podem ser considerados satisfatórios, tanto do ponto de vista da precisão dos resultados quanto na rapidez de obtenção dos mesmos. Ainda assim, sugere-se algumas melhorias e funcionalidades que podem ser implementadas com o objetivo de que o software apresentado se aplique melhor à solução de problemas de reconhecimento de padrões, dentre eles:

- Realizar experimentos com outras bases de dados;

- Acrescentar outras redes da família ART ao software;
- Implementar uma funcionalidade que permita ao usuário visualizar as matrizes de confusão para cada experimento;
- Permitir ao usuário a visualização do número de classes gerada;
- Permitir que o usuário combine dois ou mais modelos disponíveis no software em busca do melhor resultado.

Referências Bibliográficas

Amin, K., Sutariya, A.. *An Improvement in K-means Clustering Algorithm*. International Journal of Engineering Research & Technology. Vol. 2, Nº1, 2013.

Amorim, D. G.. *Redes ART com Categorias Internas de Geometria Irregular*. Tese de Doutorado. Santiago de Compostela, 2006.

Anderson, C.. *CS545: Assignment 4*. Disponível em <http://www.cs.colostate.edu/~anderson/cs545/assignments/solutionsGoodExamples/assignment4Barriault.pdf>. Acesso em 15/01/2013.

Bartfai, G.. *An Inprovised Learning Algorithm for the Fuzzy ARTMAP Neural Network.* IEEE, Artificial Neural Networks and Expert System, Proceedings, Nova Zelândia, pp. 34-37, 1995.

Beckmann, M.. *Algoritmos Genéticos como Estratégia de Pré-Processamento em Conjuntos de Dados Desbalanceados*. Dissertação de Mestrado, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2010.

Bishop, C.. *Neural Networks for Pattern Recognition*. Oxford UP, 1995.

Carpenter, G. A., Grossberg, S.. *A Massively Parallel Architecture For A Self-Organizing Neural Pattern Recognition Machine.* Computer Vision, Graphics, And Image Processing, Vol. 37, pp. 54-115, 1987a.

Carpenter, G. A., Grossberg, S.. *Invariant Pattern Recognitin and Recall by an Attentive Self-Organizing ART Architecture in a NonstationryWorld*. IEEE, First International Conference on Neural Networks. San Diego, CA.1987b.

Carpenter, G. A., Grossberg, S., Reynolds, J. H.. *ARTMAP: Supervised Real-Learning And Classification Of Non-Stationary Data By A Self-Organizing Neural Network.* Neural Network, Vol. 4, Nº 5, pp 565-588, 1991a.

Carpenter, G. A., Grossberg, S., Rosen, D. B.. *Fuzzy ART: Fast Stable Learning And Categorization Of Analog Patterns By An Adaptive Resonance System.* Neural Network, Vol. 4, Nº 1, pp. 759-771, 1991b.

Carpenter, G. A., Grossberg, S., Markuzon, N., Reynolds, J. H. and Rosen, D. B.. *Fuzzy ARTMAP: A Neural Network Architecture for Incremental Supevised Learning of Analog Multidimentional Maps.* IEEE Transactions on Neural Networks, Vol. 6, Nº 5, pp 698-713, 1992.

Castro, A. A. M., Prado, P. P. L.. *Algoritmos Para Reconhecimento de Padrões*. Dissertação de Mestrado, Universidade de Taubaté, Taubaté, 2001.

Duda R. O., Hart, P. E.. Pattern Classification and Scene Analysis. Wiley-Interscience, USA, 1st edition, 1973.

Duda, R. O., Hart, P. E., D.G. Stork, Pattern Classification, Wiley, 2nd. Edition, 2001.

Ene, M.. *Neural Network-Based Approach to Discriminate Healthy People from those with Parkinson's Disease.* Annals of the University of Craiova. Vol. 35, pp. 112-116. 2008.

Ferreira, M. F.. *O Reconhecimento de Padrões*. Dissertação de Mestrado. Universidade de Brasília, Brasília, 1994.

Ferreira, W. P.. *Analise Dinâmica de Contingências de Sistemas de Energia Elétrica por Redes Neurais Baseadas na Teoria da Ressonância Adaptativa.* Tese de Doutorado, Unesp, 2003.

Freitas, A. R. R., Lopes, R. A., Silva, R. C. P., Guimarães, F. G.. *Classificação de Dados Através de Árvores de Perceptrons Geradas por Evolução Diferencial*. Brasilian Conference on Intelligent Systems. Curitiba, Paraná, 2012.

Gonzalez, R. C., Tou, J. T.. *Pattern Recognition Principles*. Addison-Wesley Publishing Company, Massachusetts, 1981.

Granger, E., Blaquère, Y., Savaria, Y., Cantin, M. A. and Lavoie, P. *A VLSI Architecture for Fast Clustering with Fuzzy ART Neural Networks.* International Journal of Microelectronics System Integration, Vol.5, Nº1, pp. 3-18, 1997.

Grossberg, S. *Adaptative Pattern Classification and Universal Recoding, I: Parallel Development and Coding of Neural Feature Detectors.* Biological Cybernetics, Vol. 23, Nº 3, pp. 121-134, 1976.

Hebb, D. O.. *The Organization of Behavior: A Neurophychological Theory.* Wiley. New York, 1949.

Hopfield, J. J.. *Neural Networks and Physical Systems With Emergent Collective Computational Abilities.* Proceedings of the National Academy of Sciences, USA, Vol. 79, PP 2554-2558, 1982.

Jain, A. K., Duin, R. P. W., Mao, J.. Statical *Pattern Recognition: a Review*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, Nº1, Janeiro 2000.

Isoda, L. Y. *Análise da Estabilidade Estática de Tensão de Sistemas Elétricos de Potência Usando uma Rede Neural Baseada na Teoria da Ressonância Adaptativa.* Tese de Doutorado, Unesp, 2009.

Kenaya, R.. *Euclidean ART Neural Networks.* Proceedings of the World Congress on Engineering and Computer Science 2008, WCECS2008, San Francisco, USA, 2008.

Kohonen, T.. *Self-Organization and Associative Memory.* Spinger-Verlang, New York, 3 ed., 1989.

Kohonen, T.. *Self-Organized Formation of Topologically Correct Feature Maps.* Biological Cybernetics, Vol. 43, pp 59-69, 1982.

Lopes, M. L. M.. *Desenvolvimento De Redes Neurais Para Previsão de Cargas Elétricas De Sistemas De Energia Elétrica.* Tese de Doutorado, Unesp, 2005.

Lopes, M. L. M., Lotufo, A. D. P., Marchiori, S. C., Minussi, R. , SILVEIRA, M. C. G.. *Neural Network Based on Adaptive Resonance Theory with Continuous Training for Multi-Configuration Transient Stability Analysis of Electric Power Systems.* Applied Soft Computing, Vol. 11, pp 706-715, 2011.

Marchiori, S. C., Minussi, C. R., Lotufo, A. D. P.. *Rede Neural ARTMAP Nebulosa para Análise de Estabilidade Transitória de Sistemas de Energia Elétrica.* Revista da Sociedade Brasileira de Redes Neurais – Learning and Nolinear Models, Vol. 1, Nº1, pp. 61-78, 2002.

McCulloch, W. S., Pitts, W.. *A Logical Calculus of the Ideas Immanent in Nervous Activity.* Bulletin of Mathematical Biophysics, Vol. 5, pp 115-133, 1943.

Mendes Filho, E.F. and Carvalho, A.C.P.L. *Evolutionary Design of MLP Neural Network Architectures.* IV SBRN – Simpósio Brasileiro de Redes Neurais, Campos do Jordão, São Paulo, pp. 58-65, 1997.

Minsky, M. L. and Papert, S. A.. *Perceptrons.* Cambridge, MA: MIT Press, 1969.

Mitra, S. *Fuzzy Inferencing with ART Newtwork* . IEEE International Conference on Neural Networks, Orlando, EUA. Annals. Orlando, vol. 2. pp. 1230-1234, 1994.

Moreno, A. L.. *Análise Da Estabilidade Transitória Via Rede Neural ART-ARTMAP Fuzzy Euclidiana Modificada Com Treinamento Continuado.* Tese de Doutorado, Unesp, 2010.

Mangasarian,O. L. and Wolberg, W. H.. *Cancer Diagnosis via Linea Programming*", SIAM News, Vol. 23, Nº5, pp 1-18, 1990.

Negri, T. T.. *Reconhecimento de Padrões Via Wavelets: Aplicação no Monitoramento de Estruturas a Partir do Efeito do Multicaminho Proveniente de Sinais de Satélites GPS*. Dissertação de Mestrado. Presidente Prudente, 2012.

Neto, A. R. R., Sousa, R., Barreto, G. A., Cardoso, J. S.. *Diagnostic of Pathology on the Vertebral Column with Embeded Reject Option*. 5th Iberian Conference. Las Palmas de Gran Canaria, Spain, pp 588-595, 2011.

Perelmuter, G., Carrera, E. V., Vellasco, M., Pacheco, A.. *Reconhecimento de Imagens Bidimensionais Utilizando Redes Neurais Artificiais*. IN: Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens. 7, 1995. São Carlos. Anais, pp.197-203. São Carlos. 1995.

Prado, A. Jr., Elfes, A.. *Um Projeto em Reconhecimento de Padrões de Forma*. Monografia de Graduação, ITA, São José dos Campos, SP, 1975.

Ribeiro, J. S., Teófilo, R. F., Augusto, F., Ferreira, M. M. C.. *Reconhecimento de Padrões não-supervisionados de Cafés com Diferentes Graus de Torra e Processamentos Utilizando Cromatogramas obtidos por SPME-GC-FID Alinhados pelo Algoritmo COW.* 29º Reunião Anual da Sociedade Brasileira de Química. Águas de Lindóia, 2006.

Rosenblat, F.. *Perceptron.* Massachussets Institute of Thecnology, 1958.

Salgado, R. M.. Sistema *de Suporte À Decisão Para Previsão De Carga Por Barramento*. Tese de Doutorado, Unicamp, 2009.

Schalkoff, R. J.. *Pattern Recognition: Statistical, Structural and Neural Approaches*. Wiley, New York, 1992.

Shadvar, A.. Dimension *Reducion by Mutual Information Feature Extraction*. International Journal of Computer Science & Information Technoloy. Vol. 4, Nº3, 2012.

Silva, J. A., Moreno, A. L.. *Análise dos parâmetros da Rede Neural ART Euclidiana Modificada utilizando a clusterização de pontos do plano.* Congresso Sul Mineiro de Iniciação Científica. 2012.

Steiner, M. T. A.. *Uma Metodologia para o Reconhecimento de Padrões Multivariados com Resposta Dicotômica*. Tese de Doutorado, Florianópolis, SC, 1995.

Theodoridis, S.. *Pattern Recognition*. Academic Press, USA, 1st edition, 1999.

Todesco, J. L.. *Reconhecimento de Padrões Usando Rede Neuronal Artificial com uma Função de Base Radial: uma Aplicação na classificação de Cromossomos Humanos*. Tese de Doutorado, Florianópolis, SC, 1995.

Werbos, P. J.. *Beyond Regression: New Tools for Prediction and Analysis in The Behavioral Sciences.* Master Thesis, Harvard University, 1974.

Widrow, B.. *Generalization and Information Storage in Network of Adaline neurons.* Self-Organization Systems, pp 435-461, Washington, DC: Spartan Books, 1962.

Widrow, B., Hoff, M. E.. *Adaptative Switching Circuits.* Ire Wescon Conventon Record, pp 96-104, 1960.

Zanchettin, C., Ludermir, T. B.. *Sistemas Neurais Híbridos para Reconhecimento de Padrões em Narizes Artificiais.* Revista Controle & Automação. Vol. 16, Nº 2, 2005.