

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA MECÂNICA**

**ANÁLISE DA PRESSÃO PLANTAR DA MARCHA DE  
AUTISTAS POR DINÂMICA SIMBÓLICA OTIMIZADA POR  
ALGORITMO GENÉTICO**

**LYNWOOD LIVI DE SOUZA**

**ORIENTADOR: DR. MARIO OLAVO MAGNO DE CARVALHO  
COORIENTADORA: DR.<sup>a</sup> SUÉLIA DE SIQUEIRA RODRIGUES  
FLEURY ROSA**

**TESE DE DOUTORADO EM CIÊNCIAS MECÂNICAS**

**PUBLICAÇÃO: ENM.DM-39/2016  
BRASÍLIA/DF: DEZEMBRO – 2016**

**UNIVERSIDADE DE BRASÍLIA  
FACULDADE DE TECNOLOGIA  
DEPARTAMENTO DE ENGENHARIA MECÂNICA**

**ANÁLISE DA PRESSÃO PLANTAR DA MARCHA DE  
AUTISTAS POR DINÂMICA SIMBÓLICA OTIMIZADA POR  
ALGORITMO GENÉTICO**

**LYNWOOD LIVI DE SOUZA**

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

**APROVADA POR:**

---

**PROF. MARIO OLAVO MAGNO DE CARVALHO  
UNB/ENM (ORIENTADOR)**

---

**SUÉLIA DE SIQUEIRA RODRIGUES FLEURY ROSA  
UNB / FGA (CO-ORIENTADORA)**

---

**ALINE SOUZA DE PAULA  
UNB / ENM (MEMBRO INTERNO)**

---

**TITO DIAS JÚNIOR  
SETOR TÉCNICO-CIENTÍFICO / POLÍCIA FEDERAL / RO (MEMBRO EXTERNO)**

---

**EDSON PAULO DA SILVA  
UNB / PPMEC (MEMBRO EXTERNO)**

**BRASÍLIA/DF, 02 DE DEZEMBRO DE 2016**

## **FICHA CATALOGRÁFICA**

SOUZA, LYNWOOD LIVI DE

Análise da Pressão Plantar da Marcha de Autista por Dinâmica Simbólica Otimizada por Algoritmo Genético [Distrito Federal] 2016.

117 p., 210 x 297 mm (ENM/FT/UnB, Doutor, Ciências Mecânicas, 2016).

Tese de Doutorado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Mecânica.

- |                              |                              |
|------------------------------|------------------------------|
| 1. Dinâmica Simbólica        | 2. Algoritmo Genético        |
| 3. Pressão Plantar da Marcha | 4. Redes Neurais Artificiais |
| 4. Autismo                   |                              |

I. ENM/FT/UnB

II. Título (série)

## **REFERÊNCIA BIBLIOGRÁFICA**

SOUZA, L. L. (2016). Análise da Pressão Plantar da Marcha de Autista por Dinâmica Simbólica Otimizada por Algoritmo Genético. Tese de Doutorado em Ciências Mecânicas, Publicação em ENM.DM-39/2016, Departamento de Engenharia Mecânica, Universidade de Brasília, Brasília, DF, 117 p.

## **CESSÃO DE DIREITOS**

AUTOR: Lynwood Livi de Souza.

TÍTULO: Análise da Pressão Plantar da Marcha de Autista por Dinâmica Simbólica Otimizada por Algoritmo Genético.

GRAU: Doutor ANO: 2016

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

---

Lynwood Livi de Souza

Avenida Coronel Gaspar nº 117, Centro.

76.400-000 Uruaçu – GO – Brasil.

Correio Eletrônico: lynwoodfera@gmail.com

# DEDICATÓRIA

---

Dedico este trabalho a minha mãe e também à memória eterna de meu pai. Pois juntos abdicaram de diversas coisas para me dar uma educação de qualidade. Eu sei tudo que vocês passaram por mim, tenha certeza que esta conquista não é minha, é de vocês.

Obrigado por tudo, te amo mãe, te amo pai (saudades).

# AGRADECIMENTOS

---

Primeiramente a Deus, pelas imensas bênçãos em minha vida.

À minha esposa Thiane e ao meu filho Wood, pela grandiosa felicidade de tê-los em minha vida, proporcionando carinho, amor e força na busca dos grandes objetivos.

À minha irmã e sobrinhos, por fazerem sempre parte de minha vida, e razão de minha essência.

Ao meu amigo Gilderlan Girão, pelo companheirismo e recepção sempre especial e agradável em seu lar.

Ao meu orientador Mario Olavo, pelo imenso esforço, dedicação e por acreditar neste trabalho.

À minha co-orientadora Suélia de Siqueira Rodrigues Fleury Rosa, pelo magnífico apoio, presença, disponibilidade e entusiasmo neste trabalho.

À coordenação de Pós-Graduação em Ciências Mecânicas, pelos imensos incentivos.

Ao Instituto Federal de Goiás, pelo apoio em minha qualificação.

À FAPEG / CAPES, pelo apoio financeiro em forma de bolsa.

# LISTA DE ABREVIACOES

---

AG	Algoritmo Genético
AGs	Algoritmos Genéticos
DNA	Ácido Desoxirribonucléico
DS	Dinâmica Simbólica
EMG	Eletromiograma
EEG	Eletroencefalograma
GAU	Grupo de autistas
GCO	Grupo de Controle
MLP	Perceptron Multicamada
PPM	Pressão Plantar da Marcha
PPMs	Pressões Plantares da Marcha
SNC	Sistema Nervoso Central

# LISTA DE SÍMBOLOS

---

$\mu$	Tamanho do espaço de partição
$A_i$	Probabilidades individuais da sequência $i$ do histograma $A$
$B_i$	Probabilidades individuais da sequência $i$ do histograma $B$
$f(x)$	Função de aptidão
$\ell$	Limiar de particionamento
$L$	Vetor das frequências relativas
$m$	Tamanho da sequência de símbolos
$n$	Valor da quantidade de símbolos
$N_{obs}$	Número de colunas do histograma com valor diferente de zero
$p_{i,L}$	Frequência relativa dos valores dos índices $i$
$S_i$	Série simbólica
$\alpha$	Verdadeiros Positivos
$\beta$	Falsos Negativos
$\gamma$	Verdadeiros Negativos
$\delta$	Falsos Positivos
$\varepsilon$	Seletividade Positiva (ou Valor Preditivo Positivo)
$\theta$	Função do quantificador de Precisão
$\pi$	Sensibilidade (ou Relação Verdadeira Positiva)
$\sigma$	Entropia de Shannon
$\sigma_j$	Entropia dos indivíduos do GAU
$\sigma_k$	Entropia dos indivíduos do GCO
$\tau$	Seletividade Negativa (ou Valor Preditivo Negativo)
$\varphi$	Especificidade (ou Relação Verdadeira Negativa)

# LISTA DE FIGURAS

---

FIGURA 1 - TRANSMISSÃO DO IMPULSO NERVOSO ENTRE O CÉREBRO E O PÉ. ....	38
FIGURA 2 - SISTEMA DE SENSORES COM PALMILHA. ....	39
FIGURA 3 - PLATAFORMA DE PRESSÕES. ....	40
FIGURA 4 - SIMBOLIZAÇÃO DE SÉRIE TEMPORAL BINÁRIA ( $N=2$ ). ....	43
FIGURA 5 - DISTRIBUIÇÃO DAS FREQUÊNCIAS DA SÉRIE TEMPORAL DA TABELA 1. ....	44
FIGURA 6 - TRANSFORMAÇÃO SIMBÓLICA COM $N=5$ . ....	45
FIGURA 7 – (A) PROCESSO DE SIMBOLIZAÇÃO DE SÉRIE TEMPORAL; (B) TABULAÇÃO DE UM HISTOGRAMA DA SEQUÊNCIA SIMBÓLICA. ....	47
FIGURA 8 - FLUXOGRAMA DE OTIMIZAÇÃO UTILIZANDO O ALGORITMO GENÉTICO. ....	51
FIGURA 9 - ALGORITMO GENÉTICO TÍPICO. ....	52
FIGURA 10 - INDIVÍDUOS DE UMA POPULAÇÃO E A SUA CORRESPONDENTE ROLETA DE SELEÇÃO. ....	55
FIGURA 11 - OPERADOR DE CRUZAMENTO EM UM-PONTO. ....	57
FIGURA 12 - OPERADOR DE CRUZAMENTO MULTIPONTO. ....	57
FIGURA 13 - OPERADOR DE MUTAÇÃO. ....	58
FIGURA 14 - ARQUITETURA DE REDE DE CAMADA ÚNICA. ....	63
FIGURA 15 - DIAGRAMA DE UMA REDE NEURAL ARTIFICIAL MULTICAMADA, COM DUAS CAMADAS INTERMEDIÁRIAS. ....	64
FIGURA 16 - DIAGRAMA DE DUAS REDES RECORRENTES DIFERENTES. ....	65
FIGURA 17 - ARQUITETURA DE UM MODELO MLP 4-4-4-2. ....	68
FIGURA 18 - FLUXO DE SINAIS NOS NEURÔNIOS DE UM MODELO MLP. ....	68
FIGURA 19 - DISTRIBUIÇÃO DOS SENSORES NA PLANILHA. ....	71
FIGURA 20 - FLUXOGRAMA DA OTIMIZAÇÃO DOS PARÂMETROS DA DINÂMICA SIMBÓLICA POR ALGORITMO GENÉTICO. ....	72
FIGURA 21 - FLUXOGRAMA PARA AVALIAÇÃO DO DESEMPENHO DA METODOLOGIA. ....	72
FIGURA 22 - FLUXOGRAMA DAS FASES PARA CLASSIFICAÇÃO AUTOMÁTICA DOS SINAIS DE PPM. ....	73
FIGURA 23 - RELACIONAMENTO DAS TABELAS DO BANCO DE DADOS. ....	74
FIGURA 24 - ESTRUTURA DO CROMOSSOMO CODIFICADO PARA O MÉTODO. ....	75
FIGURA 25 - CROMOSSOMO CODIFICADO EM DUAS SESSÕES COM SUA REPRESENTAÇÃO DECIMAL. ....	75
FIGURA 26 - CROSSOVER DE UM-PONTO GERANDO NOVOS INDIVÍDUOS. ....	77



FIGURA 27 - TRANSIÇÃO DE GERAÇÃO .....	78
FIGURA 28 – ILUSTRAÇÃO SINTÉTICA DO MÉTODO DA DINÂMICA SIMBÓLICA COM TRANSFORMAÇÃO SIMBÓLICA ESTÁTICA.....	82
FIGURA 29 - METODOLOGIA PARA CLASSIFICAÇÃO AUTOMÁTICA DOS SINAIS .....	85
FIGURA 30 – (A) SINAL DO ELETRODO #5 DO SUJEITO DO GAU #1 E (B) RESPECTIVA TRANSFORMAÇÃO SIMBÓLICA.....	91
FIGURA 31 – (A) SINAL DO ELETRODO #5 DO SUJEITO #1 DO GCO E (B) RESPECTIVA TRANSFORMAÇÃO SIMBÓLICA.....	91
FIGURA 32 – (A) HISTOGRAMA DA SÉRIE SIMBÓLICA DO ELETRODO #7 DO SUJEITO GAU #1. (B) HISTOGRAMA DA SÉRIE SIMBÓLICA DO ELETRODO #7 DO SUJEITO #3 DO GCO. ....	92
FIGURA 33 – DISTRIBUIÇÃO DAS MÉDIAS DAS FREQUÊNCIAS RELATIVAS DAS SÉRIES SIMBÓLICAS DOS DADOS DO GAU. ....	92
FIGURA 34 – DISTRIBUIÇÃO DAS MÉDIAS DAS FREQUÊNCIAS RELATIVAS DAS SÉRIES SIMBÓLICAS DO GCO. ....	93
FIGURA 35 – DISTRIBUIÇÃO DAS MÉDIAS DAS FREQUÊNCIAS RELATIVAS DAS SÉRIES SIMBÓLICAS ENTRE TODOS OS ELETRODOS DO GAU (A) E GCO (B). ....	94
FIGURA 36 – GRÁFICO <i>Boxplot</i> DA DISTRIBUIÇÃO DOS SINAIS POR GRUPO. ....	96
FIGURA 37 - DISTRIBUIÇÃO DOS SÍMBOLOS NOS ELETRODOS DAS SÉRIES DO GAU. ....	97
FIGURA 38 - DISTRIBUIÇÃO DOS SÍMBOLOS NOS ELETRODOS DAS SÉRIES DO GCO. ....	97
FIGURA 39 – GRÁFICO DA TABELA 18 ILUSTRANDO OS RESULTADOS DAS ESTATÍSTICAS QUANTITATIVAS. ....	102

# LISTA DE TABELAS

---

TABELA 1 – SEQUÊNCIA DE SÍMBOLOS BINÁRIOS ( $N=2$ ) DE COMPRIMENTO $M=3$ COM SEUS RESPECTIVOS CÓDIGOS DE BASE DECIMAL E FREQUÊNCIA DE OCORRÊNCIA NO SINAL MOSTRADO NA FIGURA 4. ....	44
TABELA 2 – TABELA DOS PERFIS DOS SINAIS INDICADOS NO TREINAMENTO DA RNA. ....	84
TABELA 3 – CONFIGURAÇÃO DOS OPERADORES GENÉTICOS UTILIZADOS NA OTIMIZAÇÃO.....	89
TABELA 4 – RESUMO DA OTIMIZAÇÃO DA SIMBOLIZAÇÃO DAS SÉRIES TEMPORAIS POR ALGORITMO GENÉTICO. ....	90
TABELA 5 – PORCENTAGEM MÉDIA POR SÍMBOLO EM DADOS DAS SÉRIES SIMBÓLICAS DO GAU. ....	94
TABELA 6 - PORCENTAGEM MÉDIA POR SÍMBOLO EM DADOS DAS SÉRIES SIMBÓLICAS DO GCO. ....	95
TABELA 7 - MEDIDAS DESCRITIVAS DO SINAL POR ELETRODO DAS SÉRIES DO GAU. ....	95
TABELA 8 - MEDIDAS DESCRITIVAS DO SINAL POR ELETRODO DAS SÉRIES DO GCO. ....	95
TABELA 9 - TESTE QUI-QUADRADO.....	98
TABELA 10 – COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #1.....	99
TABELA 11 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #2. ....	99
TABELA 12 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #3. ....	99
TABELA 13 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #4. ....	100
TABELA 14 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #5. ....	100
TABELA 15 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #6. ....	100
TABELA 16 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #7. ....	101
TABELA 17 - COMPARAÇÃO ESTATÍSTICA <i>T STUDENT</i> DAS FREQUÊNCIAS DE SÍMBOLOS DO ELETRODO #8. ....	101
TABELA 18 - ESTATÍSTICA QUANTITATIVA DA DISTRIBUIÇÃO DAS SÉRIES SIMBÓLICAS.....	102
TABELA 19 – PARÂMETROS PARA CONSTRUÇÃO DA RNA ALGORITMO MLP.....	103
TABELA 20 - CONJUNTO DE TREINAMENTO PARA RNA.....	104
TABELA 21 – RESULTADO DA CLASSIFICAÇÃO DOS SINAIS PELA RNA. ....	105
TABELA 22 – RESULTADOS QUANTIFICADOS DOS TESTES DA RNA. ....	105

# SUMÁRIO

---

<b>DEDICATÓRIA .....</b>	<b>4</b>
<b>AGRADECIMENTOS .....</b>	<b>5</b>
<b>LISTA DE ABREVIACÕES .....</b>	<b>6</b>
<b>LISTA DE SÍMBOLOS .....</b>	<b>7</b>
<b>LISTA DE FIGURAS.....</b>	<b>8</b>
<b>LISTA DE TABELAS .....</b>	<b>10</b>
<b>SUMÁRIO.....</b>	<b>11</b>
<b>RESUMO.....</b>	<b>14</b>
<b>ABSTRACT .....</b>	<b>16</b>
<b>1 – INTRODUÇÃO .....</b>	<b>18</b>
1.1 HIPÓTESE .....	22
1.2 MOTIVAÇÃO E RELEVÂNCIA DA CONTRIBUIÇÃO.....	23
1.3 CONTRIBUIÇÕES ESPERADAS .....	24
1.4 OBJETIVOS .....	25
1.4.1 <i>Objetivo Geral</i> .....	25
1.4.2 <i>Objetivos Específicos</i> .....	25
1.5 ESTRUTURA DA TESE.....	26
<b>2 – AUTISMO .....</b>	<b>27</b>
2.1 INTRODUÇÃO AO AUTISMO.....	27
2.2 DIAGNÓSTICO .....	29
2.3 TRATAMENTO .....	29
2.4 BIOMARCADORES.....	30

2.5	CARACTERÍSTICAS MOTORAS DO AUTISMO .....	32
<b>3</b>	<b>– PRESSÃO PLANTAR DA MARCHA .....</b>	<b>35</b>
3.1	CONCEITOS GERAIS.....	35
3.2	RECEPTORES SENSITIVOS CUTÂNEOS PLANTARES.....	37
3.3	SISTEMAS DE MEDIÇÃO.....	39
<b>4</b>	<b>– DINÂMICA SIMBÓLICA .....</b>	<b>42</b>
4.1	DEFINIÇÕES GERAIS .....	42
4.2	DEFINIÇÃO DE SÍMBOLOS .....	45
4.3	CODIFICAÇÃO .....	46
<b>5</b>	<b>– ALGORITMO GENÉTICO.....</b>	<b>48</b>
5.1	INTRODUÇÃO AO ALGORITMO GENÉTICO .....	48
5.2	ELEMENTOS DO ALGORITMO GENÉTICO .....	52
5.3	REPRESENTAÇÃO E CODIFICAÇÃO.....	53
5.4	POPULAÇÃO .....	53
5.5	SELEÇÃO .....	54
5.6	FUNÇÃO DE APTIDÃO .....	56
5.7	CRUZAMENTO .....	56
5.8	MUTAÇÃO.....	57
5.9	PARÂMETROS GENÉTICOS.....	59
5.10	APLICAÇÃO.....	60
<b>6</b>	<b>– REDES NEURAIS ARTIFICIAIS PARA CLASSIFICAÇÃO .....</b>	<b>61</b>
6.1	INTRODUÇÃO ÀS REDES NEURAIS ARTIFICIAIS .....	61
6.2	ARQUITETURA DE RNA .....	62
6.3	APRENDIZADO.....	65
6.4	TREINAMENTO.....	66
6.5	PERCEPTRON DE MÚLTIPLAS CAMADAS - MLP .....	67
<b>7</b>	<b>– METODOLOGIA.....</b>	<b>70</b>
7.1	DADOS DA PRESSÃO PLANTAR DA MARCHA.....	70
7.2	FLUXOGRAMAS DAS ETAPAS DA METODOLOGIA .....	71
7.2.1	<i>Otimização da Dinâmica Simbólica por Algoritmo genético .....</i>	<i>72</i>
7.2.2	<i>Avaliação Do Desempenho Da Metodologia Desenvolvida.....</i>	<i>72</i>
7.2.3	<i>Identificação automática da presença de autismo pelas séries de pressão plantar da marcha por meio de redes neurais artificiais .....</i>	<i>73</i>
7.3	ESTRUTURAÇÃO DOS DADOS .....	73
7.4	OTIMIZAÇÃO DOS PARÂMETROS DA DINÂMICA SIMBÓLICA POR ALGORITMO GENÉTICO .....	74
7.4.1	<i>Codificação dos indivíduos e população inicial.....</i>	<i>74</i>

7.4.2	<i>Função de Aptidão (avaliação)</i> .....	75
7.4.3	<i>Seleção dos indivíduos</i> .....	77
7.4.4	<i>Operadores genéticos</i> .....	77
7.4.5	<i>Critério de Parada</i> .....	78
7.5	SIMBOLIZAÇÃO DAS SÉRIES TEMPORAIS .....	79
7.6	IDENTIFICAÇÃO AUTOMÁTICA DA PRESENÇA DE AUTISMO .....	82
7.7	ANÁLISE ESTATÍSTICA DAS SÉRIES SIMBÓLICAS .....	85
7.8	ANÁLISE ESTATÍSTICA DO DESEMPENHO DA RNA .....	87
<b>8</b>	<b>– RESULTADOS E DISCUSSÕES .....</b>	<b>89</b>
8.1	OTIMIZAÇÃO DE PARÂMETROS.....	89
8.2	ANÁLISE DA PPM POR DINÂMICA SIMBÓLICA .....	90
8.3	ANÁLISE DA DINÂMICA SIMBÓLICA POR MEIO DE TESTES ESTATÍSTICOS .....	94
8.4	CLASSIFICAÇÃO AUTOMÁTICA DAS SÉRIES SIMBÓLICAS POR REDE NEURAL ARTIFICIAL .....	103
8.5	ANÁLISE ESTATÍSTICA DA CLASSIFICAÇÃO POR RNA .....	105
<b>9</b>	<b>– CONCLUSÃO.....</b>	<b>106</b>
<b>10</b>	<b>– PERSPECTIVAS.....</b>	<b>109</b>
<b>11</b>	<b>– REFERÊNCIAS .....</b>	<b>110</b>

# RESUMO

---

O autismo infantil é um transtorno de início precoce que afeta predominantemente três habilidades: interação social, comunicação/linguagem e comportamento. As consequências desse transtorno são diversas. Várias teorias foram criadas na tentativa de explicar as características presentes em pessoas com autismo, contudo a etiologia permanece desconhecida. Uma grande particularidade desse distúrbio é a complexidade de seu diagnóstico conclusivo. Para uma completa avaliação, são necessários diversos profissionais de diferentes áreas – como pediatras, psicólogos e neurologistas –, e isso leva a uma peregrinação dos pais até que se conclua o diagnóstico. A falta de um biomarcador é bastante discutida na literatura científica, pois diminuiria o tempo de diagnóstico e, conseqüentemente, anteciparia as terapias indicadas aos indivíduos portadores do autismo. Tais terapias são importantes, pois possibilitam aos indivíduos o desenvolvimento necessário para se socializar e conseguir ter uma vida normal. Com base na discussão da literatura científica, existem défices motores nos autistas, os quais, frequentemente são refletidos na dinâmica da marcha. Este trabalho tem como objetivo analisar, por Dinâmica Simbólica (DS) otimizada por Algoritmo Genético (AG) os dados de Pressão Plantar da Marcha (PPM) de um grupo de autistas (GAU), a fim de encontrar estruturas que possam propor a PPM como um biomarcador importante no diagnóstico precoce do autismo. Com base na etiologia mecânica da passada, busca-se caracterizar como a dinâmica da pressão da caminhada de um autista é diferente em certas regiões do pé. A metodologia desenvolvida neste trabalho iniciou-se com a obtenção dos dados de PPM do GAU e grupo de controle (GCO), desenvolveu-se um Algoritmo Genético (AG) para otimizar os parâmetros da DS. Para tanto, realizou-se a simbolização das séries temporais por DS, e, em seguida, analisaram-se estatisticamente os resultados, os quais mostraram a existência

de alteração significativa na distribuição, indicando evidências de estruturas determinísticas nas amostras de PPM, o que permitiu distinguir características dos autistas em sua pressão plantar ao caminhar. Desta forma demonstra-se que o método da Dinâmica Simbólica otimizada por Algoritmo Genético para análise dos dados de PPM é eficiente em detectar alterações na pressão plantar da marcha de portadores de autismo.

**Palavras Chaves:** Dinâmica Simbólica, Algoritmo Genético, Pressão Plantar da Marcha, Redes Neurais Artificiais e Autismo.

# ABSTRACT

---

Childhood autism is a disorder of early start which predominantly affects three abilities: social interaction, communication/language and behavior. The consequences of such disorder are diverse. Several theories have been created in the attempt to explain the characteristics which are present in people with autism, although the etiology remains unknown. A great particularity of that disorder is the complexity of its conclusive diagnosis. For a complete evaluation, several professionals of different areas are needed – such as pediatricians, psychologists and neurologists –, and that leads to a pilgrimage of the parents until the diagnosis is concluded. The lack of a biomarker is broadly discussed in scientific literature, because it would reduce the time of the diagnosis and, consequently, it would anticipate the indicated therapies to the individuals with autism. Those therapies are important, because they allow the individuals to have necessary development to socialize and manage to have a normal life. Based on the discussion of scientific literature, several authors cited here confirm motor deficits in autistic people who are frequently reflected on the dynamics of their gait. This study aims to analyze, through Symbolic Dynamics (DS) optimized by Genetic Algorithm (AG), the data from Gait Plantar Pressure (PPM) of a group of autistic people (GAU), in order to find structures which propose a PPM as an important biomarker in early diagnosis of autism. Based on the mechanical etiology of the pace, we seek to characterize how the dynamics of gait pressure of an autistic person is different in certain regions of the foot. The methodology developed in this study started with the data gathering of the PPM from the GAU and a control group (GCO), and a Genetic Algorithm (AG) has been developed to optimize the parameters of the DS. In order to do so, a symbolization of the temporal series through DS was carried out, and, subsequently, the results were analyzed statistically, which showed the existence of a significant alteration in the distribution, indicating evidence of deterministic structures



in samples of PPM, allowing the distinction of the characteristics of autistic people in their plantar pressure while walking, demonstrating that the method of Symbolic Dynamics optimized by Genetic Algorithm to analyze the data of PPM is efficient in the detection of change gait plantar pressure of autist.

**Keywords:** Symbolic Dynamics, Genetic Algorithm, Gait Plantar Pressure, Artificial Neural Networks, Autism.

# 1 – INTRODUÇÃO

---

O grande desafio da medicina nos tempos atuais é a busca eminente da cura de doenças até então incuráveis e a melhora das terapias para os pacientes acometidos por tais doenças. Sabe-se que várias doenças não são propriamente causadoras de mortes, mas afetam a qualidade de vida das pessoas por elas acometidas, e conseqüentemente dos familiares, que necessitam mudar seus hábitos e passam a conviver com uma vida adaptada e cheia de mudanças em seus costumes.

Os tabus e os preconceitos gerados pela sociedade aos pacientes com algum tipo de doença incurável causam o afastamento do convívio social, o que aumenta a dificuldade de terem uma vida normal por meio do convívio e do contato humano. Quando essa doença ou transtorno tem como principal sintoma o isolamento social, leva ao distanciamento maior por parte da sociedade quanto ao entendimento e compreensão da moléstia. O transtorno em que na maioria dos casos o indivíduo busca o isolamento social é chamado de Autismo ou Síndrome do Aspecto Autista.

A etiologia do autismo é complexa e seu sistema patológico é desconhecido. É um distúrbio heterogêneo com seu diagnóstico determinado com base em critérios. Muitas pesquisas indicam que a base do autismo está na variedade de fatores genéticos, além de estudos considerarem como fatores causadores da síndrome as condições ambientais, neurobiológicas, neuroanatômicas, metabólicas e imunológicas (REGO, 2012).

Dada a complexidade desse distúrbio, vêm sendo estudado exaustivamente biomarcadores, com a esperança de garantir o diagnóstico e tratamento precoce para alcançar uma abordagem terapêutica mais eficaz (KINDREGAN *et al.* (2015); ESPOSITO *et al.* (2011); PEREIRA *et al.* (2014) e ZWAIGENBAUM *et al.* (2015)).

A identificação precoce do autismo é essencial para garantir que crianças possam obter intervenções especializadas, baseadas em evidências, que venham ajudar a otimizar os resultados de longo prazo (JESTE & GESCHWIND, 2014). A identificação precoce também ajuda a reduzir a "odisséia de diagnóstico", um fato estressante que as famílias experimentam até sua conclusão, pois, em inúmeros casos, exige-se a participação de vários profissionais com diversos laudos e análise para, enfim, diagnosticar o autismo (PERERA *et al.*, 2014).

No Brasil, existe um déficit de profissionais especializados no diagnóstico desse transtorno, levando a um tempo maior de diagnóstico (VICTORIA, 2012). Nos Estados Unidos, e em muitos outros países, a média de idade das crianças diagnosticadas tem sido de 3 a 4 anos (VICTORIA, 2012); porém, mesmo assim, muitos pais já começam a notar que existe algum problema com seus filhos antes do segundo ano de vida e, em alguns casos, até no primeiro ano de vida. No Brasil, muitas crianças ainda permanecem com um diagnóstico sem resultado preciso até a idade de 6 ou 7 anos, e até mesmo por mais tempo (PERERA *et al.*, 2014).

Em alguns casos o diagnóstico tardio do autismo faz com que o indivíduo apresente problemas no padrão motor da marcha, utilizando a ponta do pé, postura incorreta do braço durante a caminhada e uma diferenciação do movimento em geral. Esses problemas podem estar relacionados diretamente com o cérebro por necessitar desses movimentos assimétricos na busca por uma melhor coordenação motora do equilíbrio (KINDREGAN *et al.* 2015; ESPOSITO *et al.* 2011). Dores, fadiga e estresse das articulações são algumas consequências desenvolvidas devido às alterações dos padrões de marcha nos autistas, afetando, assim, suas capacidades cinéticas funcionais. Problemas como esses destacam a importância do diagnóstico precoce de autismo, pois leva a uma melhora nas condições de vida do indivíduo (VICTORIA, 2012).

A anomalia no caminhar do autista traz características importantes de serem observadas e analisadas, pois a marcha humana pode ser definida como uma série de movimentos alternados rítmicos nas extremidades inferiores e do tronco, o qual permite o movimento do corpo por meio da ação coordenada de cada um dos componentes do sistema locomotor humano. Ela pode ser afetada por um grande conjunto de lesões e doenças que alteram significativamente as dinâmicas naturais de movimento. (PERERA *et al.*, 2014).

Os problemas apresentados no comportamento motor podem ser diagnosticados anteriormente à aparição das desordens sociais e anormalidades linguísticas. Diversos trabalhos analisam o desempenho motor de autistas como forma de buscar metodologias terapêuticas para melhoria de sua qualidade de vida, pois conclui-se que crianças e adultos mostram problemas de equilíbrio, movimentos repetitivos, padronizados e estereótipos motores, evidenciando a hipótese de que o desenvolvimento motor tem sido frequentemente indicado como um biomarcador precoce de autismo (ESPOSITO *et al.*, 2011; ZWAIGENBAUM *et al.*, 2015).

Kindregan *et al.* (2015), Esposito *et al.* (2011), Perera *et al.* (2014) e Zwaigenbaum *et al.* (2015) destacaram trabalhos com estudos experimentais do movimento e da marcha em autistas, que em sua maioria observaram e comprovaram diferenças nas características motoras dos padrões de marcha, como: redução do comprimento de suas passadas para garantir uma boa estabilidade, reduzida amplitude de movimento nos tornozelos e joelhos durante a marcha, e aumento da flexão do quadril.

Os estudos de Weiss *et al.* (2013) e Victoria *et al.* (2012) tiveram como objetivo examinar as diferenças na simetria da marcha entre um grupo de crianças com autismo versus grupo de controles, e os resultados sugerem que as crianças com autismo demonstram padrões simétricos típicos da marcha.

Na análise da marcha realizada por Nobile *et al.* (2011), os autores tiveram a percepção de uma marcha caracterizada mais dura com alterações posturais do tronco e dificuldades muito significativas para manter uma linha reta. Os dados obtidos sugerem a presença de uma deficiência generalizada nos movimentos, envolvendo não apenas as habilidades motoras básicas (parâmetros lineares da marcha), mas também de controle motor, o qual abrange o processo de integração de informações e processamento.

O trabalho de Lim *et al.* (2016) objetivou investigar o padrão de marcha de crianças com autismo, usando um sistema de análise de marcha. Os resultados mostraram uma redução na cadência, velocidade da marcha, comprimento do passo, e um aumento na largura do passo em crianças com autismo. Variáveis de pressão plantar destacaram as diferenças entre as áreas de pressão ativas, especialmente no retropé de crianças com autismo, concluindo que elas têm uma marcha anormal em comparação com as do grupo

de controle e, portanto, elas precisam de atenção extra para corrigir esses padrões anormais da marcha.

Os défices do desenvolvimento motor e da estabilidade do movimento devem ser observados e definidos como consequência de problemas neurológicos, sendo uma resposta principal e evidente do sintoma do autismo. Assim, deixa de ser apenas uma função adicional na observação do distúrbio e se tornando uma hipótese indicativa de que os problemas motores podem ser mais bem avaliados que outros sistemas de função cognitiva. Isso faz com que o movimento e a marcha sejam marcadores confiáveis para o diagnóstico do autismo, haja vista que, nos últimos anos, tornou-se claro que as crianças com autismo têm dificuldades com a função e coordenação motora, e esses fatores influenciam diretamente na marcha (KINDREGAN *et al.*, 2015).

A marcha sofre influência de um conjunto de fatores que podem iniciar desde o cérebro até os membros inferiores, e essa conjunção reflete diretamente na pressão plantar exercida para que haja o controle do equilíbrio correto. Isso ocorre pela existência de um conjunto disperso de receptores sensitivos cutâneos na região plantar do pé, com grande sensibilidade às pressões de contato e às variações existentes nas distribuições dessas mesmas pressões. Em contato com a superfície, os nervos existentes levam o sinal até o Sistema Nervoso Central (SNC). O SNC utiliza-se dessas informações e produz os padrões para a manutenção da postura e locomoção por meio dos músculos e, de acordo com essas trocas de informações, geram o sistema de equilíbrio corporal (TÁBUAS, 2012).

A literatura destaca os défices e características da marcha dos autistas, porém não propõem um sistema de biomarcador definido que realize o diagnóstico utilizando esses padrões da marcha. Portanto, propomos a utilização da Pressão Plantar da Marcha (PPM) como um biomarcador do autismo, através da qual conseguimos caracterizar a presença de autismo.

Neste estudo, analisamos os dados de PPM utilizando a dinâmica não linear simbólica para diferenciar os tipos de sinais da marcha, baseando-se em suas medidas que descrevem o fenômeno em estudo. A simbolização envolve a transformação de séries temporais medidas em séries de símbolos discretos, que são processados para extrair informações sobre o processo subjacente. Esse método tem sido utilizado com sucesso na caracterização e no reconhecimento de padrões temporais em vários tipos de sistemas

experimentais e sinais biológicos (ROCHA, 2013; BAUMERT *et al.*, 2015; COSTA *et al.*, 2014; SOUZA, 2012).

Na proposta de análise dos dados de PPM por Dinâmica Simbólica (DS), inserimos a aplicação de uma ferramenta de otimização para seus parâmetros. Essa inserção na metodologia inova na aplicação das análises por séries temporais, pois diminui o risco de se realizar experimentos de forma arbitrária e gerar resultados por meio de inúmeros testes experimentais com conclusões subjetivas.

As análises das PPMs podem ser realizadas com diversas técnicas que relacionam tarefas ou condições experimentais com os sinais resultantes, e posterior análise estatística. A técnica da DS, cujo uso na análise dos dados de PPM é proposto neste trabalho, é uma técnica alternativa e inovadora para análise de sinais ou séries. Ela envolve a codificação dos sinais como sequências de símbolos de um alfabeto (por exemplo, o alfabeto binário {0,1}), e a avaliação dessas sequências resultantes. Com as séries assim formadas, é possível avaliar as taxas de ocorrência de determinadas subestruturas (denominadas “palavras”), bem como a tendência de ocorrência privilegiada de algumas delas.

A proposta da dinâmica simbólica é codificar o sinal em uma sequência de símbolos, simplificando-se o processamento e a quantidade de variáveis analisadas (SOUZA, 2012; ROCHA, 2013). Com esta codificação há perda de detalhes do sinal, porém a informação robusta é preservada com o mínimo de tempo e processamento necessários (COSTA *et al.*, 2015). A dinâmica simbólica tem sido utilizada no processamento de sinais biológicos: por exemplo, em técnicas de reconhecimento de padrões para a classificação de sinais fisiológicos (ROCHA, 2013), em padrões respiratórios, em dados de ressonância magnética, na avaliação cardiovascular, em interações cardiorrespiratórias (BAUMERT *et al.*, 2015), na análise de temperatura (COSTA *et al.*, 2014) e na detecção de eventos epiléticos em sinais de EEG (SOUZA, 2012).

## **1.1 HIPÓTESE**

A hipótese assumida é de que o modelo de simbolização de séries temporais com a otimização de seus parâmetros por algoritmo genético seja suficiente para encontrar

estruturas determinísticas nos dados de Pressão Plantar da Marcha de autistas que permitam distingui-las dos dados de Pressão Plantar da Marcha de não autistas.

## **1.2 MOTIVAÇÃO E RELEVÂNCIA DA CONTRIBUIÇÃO**

O autismo faz parte de um grupo de patologias do desenvolvimento neurológico caracterizado por uma forte base genética, sendo considerado o distúrbio mais debilitante e comum nas crianças (JESTE & GESCHWIND, 2014). As deficiências na socialização, comunicação, interesses repetitivos e padrões de comportamento são características principais dos autistas e que têm um impacto sobre a maioria dos domínios de funcionamento cognitivo e adaptativo. O número de casos de autismo no mundo está estimado em 1 a cada 68 indivíduos (Centro de Controle e Prevenção de Doenças, 2014), que é 37 vezes maior do que o relatado em estudos publicados trinta anos atrás.

Mesmo com a caracterização da síndrome do autismo nas últimas décadas, técnicas de diagnósticos confiáveis e consistentes ainda não foram identificadas. Portanto, o diagnóstico dessa síndrome permanece atualmente com base nos sintomas e comportamentos do sujeito, geralmente após prejuízos sociais, comunicativos e com a presença de comportamentos estereotipados. Atualmente, a avaliação da criança é realizada por clínicos especializados, incluindo profissionais da pediatria, psiquiatria e psicologia, quando são feitos inúmeros testes que envolvem a observação do paciente e ou entrevista com o cuidador. No entanto, esses métodos podem não ser confiáveis, sendo relativamente baixos os níveis de especificidade diagnóstica em crianças. Portanto, há uma necessidade urgente de desenvolver métodos mais objetivos de detecção do autismo, principalmente biomarcadores de diagnóstico para a doença (RETICO *et al.*, 2014).

Recentemente, a comunidade científica revisou a opinião sobre o resultado e as consequências do autismo. Pesquisadores e médicos estão começando a considerá-lo como um distúrbio tratável, com base na descoberta de que alguns pacientes, graças a um diagnóstico precoce e uma intervenção intensiva, podem chegar à função cognitiva normal, deixando de ter um diagnóstico positivo da síndrome (FEIN *et al.*, 2013). Dessa forma, a investigação sobre a identificação precoce do autismo representa uma prioridade de saúde pública. Além disso, um diagnóstico e tratamento precoce poderiam ter um impacto econômico positivo, tornando possível reduzir os custos de tratamentos durante a vida de uma pessoa com autismo.

Pesquisas como as de Bolton *et al.* (2012), Zwaigenbaum *et al.* (2013), Esposito *et al.* (2011; 2009; 2008), Retico *et al.* (2014) e Flanagan *et al.* (2012) detectaram, em seus experimentos, diferenças qualitativas importantes no funcionamento motor, indicando que muitas crianças que possuíam a síndrome do autismo tinham dificuldades significativas no equilíbrio, estabilidade postural e movimentos. Destacaram também que bebês que possuíam alterações motoras nos primeiros 6-9 meses – incluindo atrasos nos reflexos protetores, assimetria, posição anormal do corpo ao rastejar, além de atrasos na maturidade dos movimentos (por exemplo, controle da cabeça e coordenação dos movimentos ao caminhar) – mais tarde foram diagnosticados com autismo. Segundo (BOLTON *et al.*, 2012), as coordenações motoras são as maiores informações para um conjunto de ferramentas de triagem de risco de autismo em bebês de 6-9 meses de idade.

Essas descobertas levantam questões interessantes sobre se as alterações no sistema motor em indivíduos com autismo, expressando um subdesenvolvimento do controle postural durante a infância (ESPOSITO *et al.*, 2011). Porém, dentre essas questões citadas acima, as cognições motoras ainda não estão sendo analisadas com atenção específica para ser um potencial biomarcador de diagnóstico do autismo (ZWAIGENBAUM *et al.*, 2015).

Este trabalho tem por objetivo destacar a utilização da Pressão Plantar da Marcha nas aplicações clínicas desta pesquisa, com particular destaque da Dinâmica Simbólica na metodologia de análise de dados, técnica que tem potencial para impactar significativamente no diagnóstico precoce do autismo.

### **1.3 CONTRIBUIÇÕES ESPERADAS**

O desafio clínico do autismo contribui para o complexo desenvolvimento de uma estratégia global de detecção precoce. No entanto, é essencial desenvolver abordagens eficazes para identificar e diagnosticar crianças com autismo logo nos primeiros anos de vida. Os ganhos por meio de uma intervenção precoce podem melhorar o funcionamento adaptativo e cognitivo do indivíduo e, finalmente, reduzir consideravelmente, para a família, os custos relacionados com as terapias (ZWAIGENBAUM *et al.*, 2015).

Sabendo da necessidade de um potencial biomarcador precoce, da não existência de biomarcadores definitivos e da evidência de falhas no desenvolvimento motor dos indivíduos com autismo pode emergir, por meio desse viés, um importante biomarcador: a



análise da Pressão Plantar da Marcha. Isso é evidenciado por estudos que garantem que crianças com autismo possuem problemas no desenvolvimento motor (RETICO *et al.*, 2014) e (ZWAIGENBAUM *et al.*, 2015).

Espera-se, por meio da análise da Pressão Plantar da Marcha utilizando o modelo não linear da simbolização das séries temporais com a otimização de seus parâmetros por AG, projetar estudos iniciais mais aprofundados do uso deste possível candidato a biomarcador capaz de auxiliar no diagnóstico precoce do autismo em crianças. Seu modelo não invasivo e o custo relativamente baixo para sua aquisição podem aliar ao benefício da solução de um problema que aflige milhões de pessoas no mundo.

## **1.4 OBJETIVOS**

### **1.4.1 OBJETIVO GERAL**

O objetivo do trabalho é avaliar os sinais de Pressão Plantar da Marcha (PPM) de um grupo de autistas (GAU) e de um grupo de controle (GCO) através da simbolização de séries temporais otimizadas por algoritmo genético, e verificar, com as informações extraídas pela metodologia, a possibilidade de distinguir, pelo sinal de PPM, quando um determinado indivíduo pertence ao grupo autista.

### **1.4.2 OBJETIVOS ESPECÍFICOS**

- Estruturação dos dados experimentais em uma base de dados consistente para o processamento, testes e análise de resultados;
- Implementação em um Algoritmo Genético uma solução capaz de otimizar os parâmetros da Dinâmica Simbólica;
- Aplicação dos parâmetros otimizados na simbolização de séries temporais dos dados de PPM;
- Análise estatística dos resultados da simbolização das séries temporais dos dados de PPM do GAU e PPM do GCO;
- Implementação dos resultados da simbolização das séries temporais em uma Rede Neural Artificial, afim de que se realize a classificação automática dos perfis de cada padrão dos sinais de PPM, gerados pela simbolização das séries temporais;

- Análise estatística da capacidade da Rede Neural Artificial em detectar automaticamente as características de cada sinal de PPM.

## **1.5 ESTRUTURA DA TESE**

A fim de contextualizar nossa pesquisa e caracterizar melhor o trabalho em questão, descrevemos, do Capítulo 2 ao Capítulo 6, as revisões bibliográficas e conceitos teóricos que delinearão as contribuições deste trabalho. Assim realizaram-se revisões sobre o autismo como patologia, as análises de pressões plantares de crianças, a dinâmica simbólica de sinais, os métodos de otimização e as redes neurais artificiais. No Capítulo 7 identificam-se os métodos específicos configurados para esta pesquisa. O Capítulo 8 reporta-se às discussões e aos resultados obtidos com a análise dos dados de Pressão Plantar da Marcha por meio da simbolização das séries temporais. O Capítulo 9 descreve as conclusões do trabalho e o Capítulo 10 resume as perspectivas da pesquisa.

## **2 – AUTISMO**

---

Neste capítulo é apresentada uma breve introdução sobre o autismo. Em seguida, descreve-se sobre os atuais tipos de diagnósticos e formas de tratamento para o transtorno, detalham-se os principais biomarcadores existentes e finaliza-se com as características motoras do autismo.

### **2.1 INTRODUÇÃO AO AUTISMO**

O autismo é uma perturbação neurobiológica global do desenvolvimento que se inicia normalmente antes dos três anos de idade e tem por características principais dificuldades na socialização, na linguagem e no comportamento. No domínio social, os sintomas incluem um défice acentuado no uso da linguagem não verbal e nos reguladores da interação social, incapacidade para resolver conflitos com pessoas da sua idade, reduzido interesse na partilha com o outro e limitada reciprocidade emocional. Na área da comunicação, verifica-se um acentuado défice ou mesmo ausência de linguagem oral, dificuldade em manter uma conversação, linguagem repetitiva e com dificuldade. Quanto aos comportamentos e interesses, estes são reduzidos e estereotipados, preocupando-se com partes ou qualidades dos objetos, não dando importância à sua funcionalidade. Também é evidente a adesão a rotinas (REGO, 2012).

O Autismo foi pela primeira vez diagnosticado por Leo Kanner (psiquiatra na Universidade Johns Hopkins), em 1943. Kanner caracterizou onze crianças com uma síndrome não descrita anteriormente, a qual designou por autismo infantil precoce. Todas as crianças estudadas revelaram incapacidade para se relacionarem com os outros e utilizarem a linguagem como veículo de comunicação e um desejo impulsivo pela manutenção do estado das coisas. Associado a esses sintomas, as crianças apresentavam

estados comórbidos de ansiedade, macrocefalia e alterações nos padrões alimentares (DANIELS *et al.*, 2014).

Mais tarde, na década de setenta, Lorna Wing realizou um estudo epidemiológico e verificou que todas as crianças com diagnóstico de autismo apresentavam uma tríade de características muito específica: (i) reduzida participação social, (ii) déficit na comunicação com lacunas na linguagem expressiva e compreensiva e (iii) dificuldades nas capacidades imaginativas e de fantasia, com comportamentos repetitivos e estereotipados (ELSABBAGH *et al.*, 2013).

Desde a publicação de Kanner ocorreram grandes avanços a nível biológico e cognitivo-comportamental do autismo. Contudo, a comunidade científica continua questionando-se acerca de sua etiologia. Apesar de ser uma perturbação neurológica, com uma provável origem pré-natal, o autismo é raramente diagnosticado antes dos dois ou três anos de idade. Assim sendo, uma vez que a maioria dos casos não é sinalizada para avaliações mais específicas antes dessa idade, ultrapassa-se a fase de desenvolvimento do potencial da criança, sendo esta, por excelência, do nascimento até aos três anos. Segundo Kandel *et al.* (2003 *apud* BRAS, 2009), o cérebro humano possui uma característica designada por plasticidade cerebral, que o torna capaz de criar novas sinapses, a qual ocorre com maior fluxo do nascimento até os seis anos. Com base nesse pressuposto, a comunidade científica procura evidências no perfil de desenvolvimento da criança que permitam estabelecer um diagnóstico do autismo o mais precoce possível (RETICO *et al.*, 2014); ZWAIGENBAUM *et al.*, 2015); ZWAIGENBAUM *et al.*, 2013).

Por conseguinte, vários autores investigam as alterações motoras em idades precoces (desde o nascimento), uma vez que, no desenvolvimento global da criança, são estas as primeiras a serem evidenciadas (RETICO *et al.*, 2014; ZWAIGENBAUM *et al.*, 2015).

Estudos retrospectivos de Zwaigenbaum *et al.* (2015) sugeriram que as crianças com autismo apresentam características motoras diferentes dos padrões normais de desenvolvimento desde o nascimento, destacando que as anomalias do movimento seriam os primeiros comportamentos a serem avaliados nos indivíduos autistas. Por sua vez, Baraneck (1999 *apud* CLIFFORD *et al.*, 2013) comparou o desenvolvimento motor de crianças com autismo e crianças ditas normais, entre os nove e os doze meses. Esse estudo

considerou as funções sensório-motoras como fortes indicadores do diagnóstico do autismo.

## **2.2 DIAGNÓSTICO**

O autismo não é uma doença específica, mas sim um distúrbio do desenvolvimento que apresenta uma forte base genética. Dada a sua heterogeneidade e grande variabilidade de sintomas, o autismo é inserido num espectro de doença (RETICO *et al.*, 2014).

O pediatra é um dos primeiros responsáveis em suspeitar do diagnóstico de autismo para implementar a intervenção mais apropriada. Deve confiar nos relatos dos pais, ter um bom julgamento clínico e habilidade para reconhecer os critérios de comportamentos que definem o autismo. Quando se suspeita do diagnóstico com base em sinais e sintomas clínicos e nos testes e escalas de rastreio, uma avaliação clínica com uma equipe multidisciplinar deve ser feita, de modo a definir o diagnóstico definitivo e procurar a sua etiologia (RAZNAHAN *et al.*, 2013; DANIELS *et al.*, 2014).

## **2.3 TRATAMENTO**

Ao contrário da maioria das doenças para as quais existe um tratamento específico definido, não existem diretrizes de tratamento protocoladas para o autismo. As estratégias adotadas visam melhorar o *status* funcional da criança, envolvendo-a num programa de intervenção integrado que promova a comunicação, a socialização e as habilidades comportamentais, adaptativas e académicas e que reduza os comportamentos indesejados, como são os comportamentos agressivos e estereotipados (REGO, 2012).

Assim sendo, as estratégias incluem a educação e suporte parental, sendo crucial a adequada informação para conhecimento dos pais. Uma intervenção precoce, por meio de métodos comportamentais, de comunicação, terapias físicas e ocupacionais, e também intervenções sociais organizadas, deve ser bem acompanhada pelas equipas. A educação especial nas escolas para as crianças com mais de 3 anos também deve ser adotada para facilitar a aprendizagem. Terapias comportamentais psicológicas mostraram-se efetivas na adaptação das crianças e na redução dos comportamentos atípicos (JESTE & GESCHWIND, 2014).

## 2.4 BIOMARCADORES

Como referido anteriormente, o diagnóstico do autismo é clínico e baseia-se na observação comportamental e na entrevista clínica. Esse tipo de diagnóstico é vantajoso porque engloba diversas variantes do autismo, colocando a etiologia noutra plano. O défice de biomarcadores existentes é explicado pela complexidade e heterogeneidade da condição, sendo o autismo conhecido pelas múltiplas causas, variando muito no tipo e na severidade dos sintomas (REGO, 2012).

Ao longo da última década, a investigação na área da anatomia e biologia do Sistema Nervoso Central tornou evidente a existência de uma série de biomarcadores do autismo, a maioria deles focada no encéfalo (JESTE & GESCHWIND, 2014). O ideal seria que esses biomarcadores revelassem a causa do autismo e fossem clinicamente úteis para complementar e melhorar o diagnóstico comportamental e ainda ajudar na detecção precoce destes distúrbios (GRZADZINSKI *et al.*, 2013)

Abaixo estão descritos os principais biomarcadores genéticos, neuroquímicos, imagiológicos, neuropatológicos e eletrofisiológicos atualmente em estudo (ZWAIGENBAUM *et al.*, 2015):

**Genéticos:** Existe um grande número de alterações genéticas associadas ao autismo e acredita-se que a interação múltipla desses fatores contribua fortemente para sua determinação. Apesar de terem sido encontradas anomalias citogenéticas na maioria dos cromossomos dos autistas, a sua relação com a expressão e a severidade da doença ainda requer muita investigação. Assim, apenas algumas alterações são comuns o suficiente para serem consideradas como um marcador genético do autismo. A maior incidência da doença nos irmãos de crianças diagnosticadas com autismo evidenciou, como já explicado, a base genética dessa patologia. Contudo, sabe-se que a maioria dessas anomalias cromossômicas acontece de forma espontânea, ou seja, surge de novo no indivíduo autista. Isso significa que a mutação ocorre pela primeira vez na família como resultado de uma alteração do DNA na célula germinativa ou no óvulo fertilizado. Contudo, ainda é incerta a sua significância funcional (FEIN *et al.*, 2013).

**Neuroquímicos:** De acordo com (DANIELS *et al.*, 2014), estudos mostram um aumento da serotonina plaquetária em 30 a 50% das crianças autistas. A serotonina é um neurotransmissor implicado em várias funções mentais, como o comportamento, sono, agressividade, ansiedade e afeto. A maior evidência que indica que a serotonina está relacionada ao autismo é a resposta a medicamentos que inibem o transporte desta substância, reduzindo condutas agressivas e estereotípias. Há evidências que apontam o autismo como uma disfunção cerebral orgânica, entre elas o retardo mental que está associado ao transtorno, na maioria dos casos, e convulsões, em alguns casos. O papel da neuroimagem para explorar e identificar a base biológica dos transtornos neurológicos ainda está no início, mas graças a ela os estudos vêm avançando. Contudo, ainda há uma inconsistência dos dados sobre a localização das alterações cerebrais, devido à grande diversidade da capacidade intelectual encontrada em autistas (RETICO *et al.*, 2014).

**Envolvendo GABA e Glutamato:** Os sistemas de neurotransmissores GABAérgico e glutamatérgico estão envolvidos na fisiopatologia do autismo. Os neurotransmissores são os pertencentes à classe dos aminoácidos neuroativos, que são o ácido gama-butírico (GABA), principal aminoácido com ação inibitória, e o glutamato, principal aminoácido com ação excitatória. Esses aminoácidos produzem respostas inibitórias e excitatórias por meio de uma alteração na condução de um ou mais canais iônicos seletivos na membrana neuronal. O GABA é o principal neurotransmissor inibitório no Sistema Nervoso Central dos mamíferos. Em virtude de sua distribuição disseminada no sistema nervoso central, os receptores de GABA influenciam muitos circuitos e funções neurais. Esses neurotransmissores inibitórios e excitatórios regulam uma série de processos do comportamento, incluindo o sono, aprendizagem, memória e sensação dolorosa. Eles podem estar envolvidos em várias doenças neurológicas e psiquiátricas, dentre elas o autismo. Assim, a modulação da sinalização GABA pelo uso de medicamentos constitui um mecanismo importante para o tratamento da hiperatividade neuronal. Os agentes terapêuticos que ativam os receptores GABA também são utilizados para sedação, ansiólise, hipnose e outras condições clínicas (TSCHINKEL, 2014). No que toca ao glutamato, várias outras evidências ajudam a relacioná-lo com o autismo. Em primeiro lugar, os sintomas de déficit de glutamato mimetizam os fenótipos comportamentais do autismo. Em segundo lugar, o receptor 2A da serotonina quando estimulado causa comportamentos semelhantes ao autismo, possivelmente pela expressão nos neurônios glutamatérgicos. Por último, sabe-se que uma diminuição do funcionamento

do GABA e o aumento da excitação pelo glutamato reduzem o limiar para desenvolvimento de convulsões, comprovando a estreita relação entre autismo e epilepsia (TSCHINKEL, 2014; RIBEIRO *et al.*, 2013).

**Neuropatológicos:** A investigação da patologia do sistema nervoso no autismo vem sendo intensificado como resultado do aumento e melhoria das técnicas neuroestruturais funcionais e histopatologia (RIBEIRO *et al.*, 2013). As técnicas de estudo *postmortem* do encéfalo humano constituem uma importante ferramenta para entender a base neurobiológica das alterações neuroanatômicas nesses distúrbios. Ainda assim, esse tipo de estudo encontra-se ainda numa fase precoce.

**Envolvendo Imagiologia:** A Ressonância Magnética é uma técnica segura para a avaliação das alterações neuroanatômicas relacionadas com o autismo. Contudo, visto que o diagnóstico só costuma ser feito por volta dos 2 a 3 anos, a informação desse exame pode ser mais indicativa dos resultados da patologia e não propriamente da causa do autismo. Estudos recentes têm utilizado a ressonância magnética funcional para demonstrar alterações nos padrões de ativação e sincronização dos circuitos corticais. Essa técnica usa o fluxo sanguíneo para quantificar a atividade neural. No encéfalo de indivíduos autistas, a conectividade funcional apresentou-se diminuída no sentido da linguagem, memória, mecanismos de resolução de problemas e processamento social (RETICO *et al.*, 2014).

**Neurofisiológicos:** No que toca a estudos eletrofisiológicos, o EEG pode constituir um exame de grande utilidade, principalmente pela elevada percentagem de indivíduos autistas com padrões encefalográficos anormais. O registro dos potenciais elétricos do cérebro tem-se mostrado útil na caracterização de funções cerebrais e na análise do estado do sistema nervoso, em pacientes com alterações sensoriais e motoras. Esse tipo de estudo poderá, no futuro, ajudar na melhor caracterização do autismo (MACCARY *et al.*, 2013).

## 2.5 CARACTERÍSTICAS MOTORAS DO AUTISMO

O movimento voluntário é a primeira manifestação de comunicação entre o bebê e os seus pais. O maior déficit observado no autismo é, sem dúvida, a falha em nível da comunicação e fraca interação com o meio, implicando, portanto, que as limitações motoras são consideradas como “sintomas associados” nos indivíduos com autismo (CHANG *et al.*, 2010). As pesquisas efetuadas na vertente do autismo incidem



principalmente nos défices de reconhecimento de expressões, limitações sociais, ignorando aparentemente o aspeto da influência corporal (MEIRELES, 2014).

Crianças com autismo apresentam várias dificuldades de controle motor, incluindo hipotonia, desenvolvimento motor global, locomotor e capacidade de controle de objetos, destreza manual, coordenação e equilíbrio, e ainda um défice geral nas respostas manuais face ao estímulo visual. Défices posturais como caminhar nas pontas dos pés ou posturas incomuns, poderão estar presentes no autismo (BEGEER *et al.*, 2013).

Existem várias evidências que sugerem que disfunções em nível do sistema motor poderão estar associadas à exibição de movimentos atípicos na perturbação do espectro autista. Essa disfunção motora poderá ser resultado das características chave no autismo (por exemplo: fraca iniciativa, dificuldades na comunicação e interação). Todavia, estudos prévios relatam a presença de movimentos anômalos a partir do primeiro ano de vida, anteriores ao surgimento das dificuldades de foro social e linguístico (NAYATE *et al.*, 2012).

Embora distúrbios motores severos não estejam necessariamente relacionados com o autismo, investigações mais recentes sugerem um maior foco nos sintomas motores, enquanto parte do processo de diagnóstico clínico é finalizado. Pode-se afirmar que os sinais motores do autismo podem ser observáveis e quantificáveis. Os sinais motores deverão servir como biomarcadores, relevantes na orientação de um diagnóstico precoce e tratamento (MEMARI *et al.*, 2013).

Note-se que uma coordenação motora empobrecida e movimentos lentos estão associados a uma participação social ineficaz e maior ansiedade durante os momentos de brincadeira na pré-escola e jardim de infância. Bhat *et al.* (2011) sugerem que as disfunções de percepção motoras presentes na infância e na vida adulta de sujeitos com autismo contribuem para as disfunções observadas em nível motor, de comunicação e social. Desse modo, a intervenção em nível motor não será apenas e exclusivamente direcionada para a melhoria de aspetos motores, como também na questão do desenvolvimento social.

As observações experimentais realizadas por Hilton *et al.* (2011) sugerem a possibilidade das limitações motoras constituírem um componente chave no fenótipo do

autismo, podendo ser um critério de inclusão para o diagnóstico. Embora as limitações motoras observadas não sejam necessariamente associadas ao autismo, já que poderão ser também observadas nas desordens de desenvolvimento da coordenação motora –, estas ocorrem comumente no autismo, estando apresentados clinicamente.

É sugerido que a disfunção neuromotora seja um sintoma base no diagnóstico do autismo, podendo ser precedida do aparecimento de disfunções em nível da comunicação e interação social (NAYATE *et al.*, 2012).

## 3 – PRESSÃO PLANTAR DA MARCHA

---

Neste capítulo são apresentados os conceitos gerais da Pressão Plantar da Marcha, explanando sua importância nas pesquisas de diagnósticos clínicos. Posteriormente, abordam-se os *Receptores Sensitivos Cutâneos Plantares*, que são responsáveis pela transmissão ao Sistema Nervoso Central de informações das pressões plantares, e finaliza-se o capítulo com os Sistemas de Medição das pressões plantares.

### 3.1 CONCEITOS GERAIS

As estruturas biomecânicas do corpo humano são constituídas de diversas características importantes para a movimentação corporal em inúmeras atividades singulares que as exercem; porém, uma delas se torna de interesse particular pela sua complexidade em atuar diretamente em contato com o meio externo – o Pé Humano, *responsável pelo equilíbrio, suporte e estabilidade durante o caminhar e marcha, dando total estabilidade ao corpo*. É uma estrutura mecânica com várias articulações, constituída por um conjunto de componentes estruturais e funcionais, tais como músculos, ossos e articulações, que interagem de forma a desempenharem uma determinada função. A sua estrutura é considerada uma das estruturas biomecânicas do corpo humano com maior complexidade e a única que atua em conjunto com uma superfície externa (PEDRO *et al.*, 2015).

De acordo com (TÁBUAS, 2012), o pé humano atua direta e constantemente com diversos fatores que envolvem a superfície plantar, que podem comprometer diretamente suas funcionalidades, desencadeando alterações da pressão plantar e, conseqüentemente, surgindo patologias e deformidades plantares.

Ao longo dos tempos o pé humano vem sendo descrito como fonte importante de informação na medida em que, por meio de um retorno (*feedback*) sensorial, permite a criação de uma interface entre o ambiente que o envolve e o sistema de equilíbrio corporal. Segundo Deepashini *et al.* (2014), esse *feedback* é estabelecido graças à presença de um conjunto disperso de receptores sensitivos cutâneos na região plantar. Estes, ao serem dotados de uma elevada sensibilidade às pressões de contato e às variações existentes na distribuição dessas mesmas pressões, possibilitam ao indivíduo, entre outras funções, o reconhecimento do ambiente e o controle dos diversos movimentos que executa em seu cotidiano.

O pé humano, ao estabelecer uma interação permanente com o meio externo, está constantemente predisposto à atuação de diferentes formas interativas que comprometem a sua funcionalidade. As referidas formas com o tempo podem induzir a produção de movimentos anormais e desencadear, conseqüentemente, alterações na distribuição da pressão em nível da superfície plantar. As alterações na distribuição da pressão são de tal forma elevadas que geram o aparecimento gradual de diversas patologias e deformidades plantares. O pé diabético, a artrite reumatoide, a doença de Charcot-Marie-Tooth e a paramiloidose são alguns dos exemplos de patologias que afetam a funcionalidade do pé (ASTRO *et al.*, 2012). É de salientar que nem sempre as áreas de maior pressão correspondem às áreas onde o paciente sente mais dor, uma vez que ele pode modificar o seu padrão normal de marcha de forma a proteger a área com dor, ou a própria presença de problemas funcionais pode ser responsável pela alteração dessa mesma distribuição (TÁBUAS, 2012; PEDRO *et al.*, 2015).

Devido à importância que desempenha, o pé tornou-se assim alvo de diversos estudos na tentativa de encontrar soluções que impeçam o aparecimento de alterações na distribuição plantar e, conseqüentemente, o comprometimento do desempenho nas diferentes atividades cotidianas do indivíduo. O avanço das pesquisas tem desenvolvido e aperfeiçoado as inúmeras técnicas que possibilitam uma análise da distribuição da pressão plantar mais aprofundada e precisa (CHIU *et al.*, 2013). Dessa forma, a análise da pressão plantar tornou-se uma ferramenta crucial na área da saúde, na medida em que especialmente permite verificar qual a sua influência da distribuição da pressão plantar no controle postural de pacientes com uma determinada patologia ou simplesmente de indivíduos que executam uma determinada atividade (THIESEN & SUMIYA, 2011).

A análise da distribuição da pressão plantar tornou-se, assim, essencial em estudos que analisam os diversos problemas motores do indivíduo (TÁBUAS, 2012). A importância do pé para o nosso corpo e as alterações por ele sofridas por inúmeros fatores passou a ser alvo de análises que envolvem patologias que possam ser diagnosticadas por meio do pé, buscando, assim, alternativas de diagnóstico que impeçam o aparecimento dessas alterações na distribuição plantar e o comprometimento de sua funcionalidade (BING *et al.*, 2016; FERNANDO *et al.*, 2016; ABRANTES *et al.*, 2011).

### **3.2 RECEPTORES SENSITIVOS CUTÂNEOS PLANTARES**

O pé humano, ao longo do tempo, vem sendo estudado e percebido como uma grande fonte de dados, que permite, por meio de um retorno sensorial, tanto pelo próprio ambiente quanto pelos movimentos executados no dia a dia obter informações importantes. Existe no pé humano um conjunto disperso de receptores sensitivos cutâneos na região plantar, com grande sensibilidade às pressões de contato e às variações existentes nas distribuições dessas mesmas pressões (FERNANDO *et al.*, 2016). Conforme observado na Figura 1, quando estamos em contato com a superfície, os nervos existentes levam o sinal até o Sistema Nervoso Central (SNC). O SNC utiliza-se dessas informações e produz os padrões para a manutenção da postura e locomoção por intermédio dos músculos e, de acordo com essas trocas de informações, geram o sistema de equilíbrio corporal (DEIRDRE *et al.*, 2015).

O pé está sempre em contato com a superfície externa e conseqüentemente traz modificações em seu trabalho natural, forçando em diversos casos mudanças na pressão plantar e na forma como ela é distribuída, podendo levar a anormalidades e complicações na vida normal da pessoa, comprometendo suas atividades (PINTO, 2014).

São diversas as pesquisas por soluções que envolvem a pressão plantar, pois o avanço nesse modelo de diagnóstico pode influenciar diretamente no desenvolvimento de tratamentos e intervenções (DEIRDRE *et al.*, 2015; LIM *et al.*, 2016; ROBERT *et al.*, 2016; QIU *et al.*, 2013).

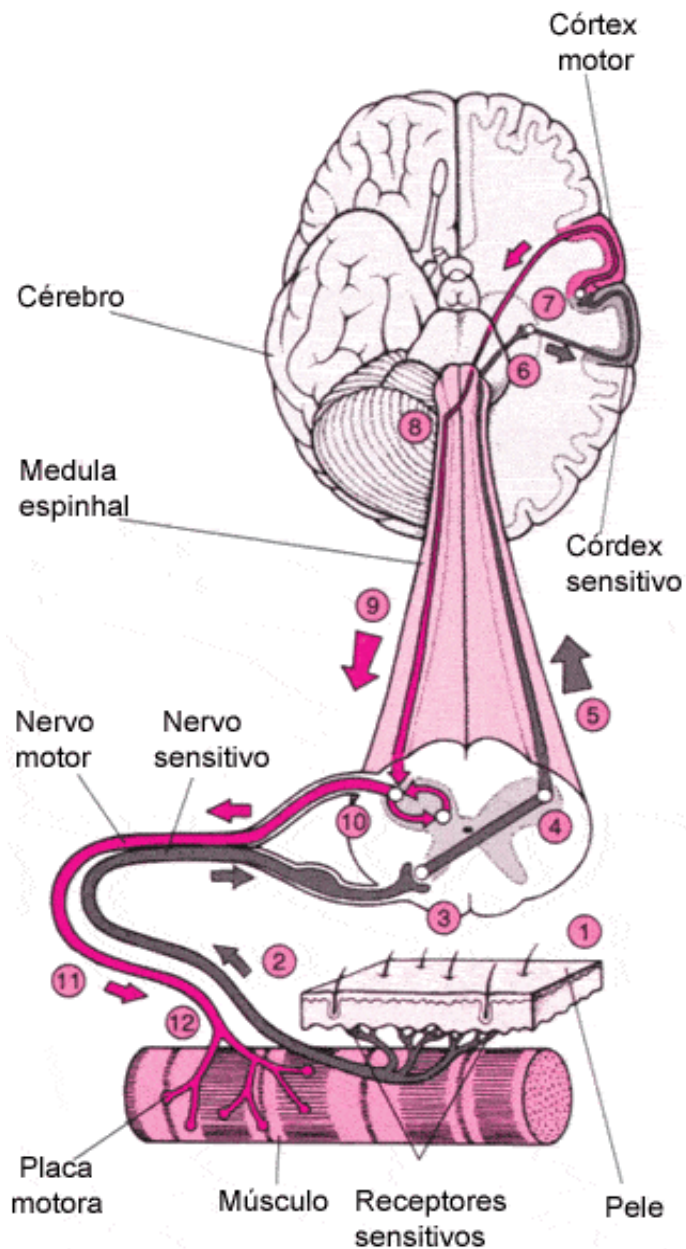


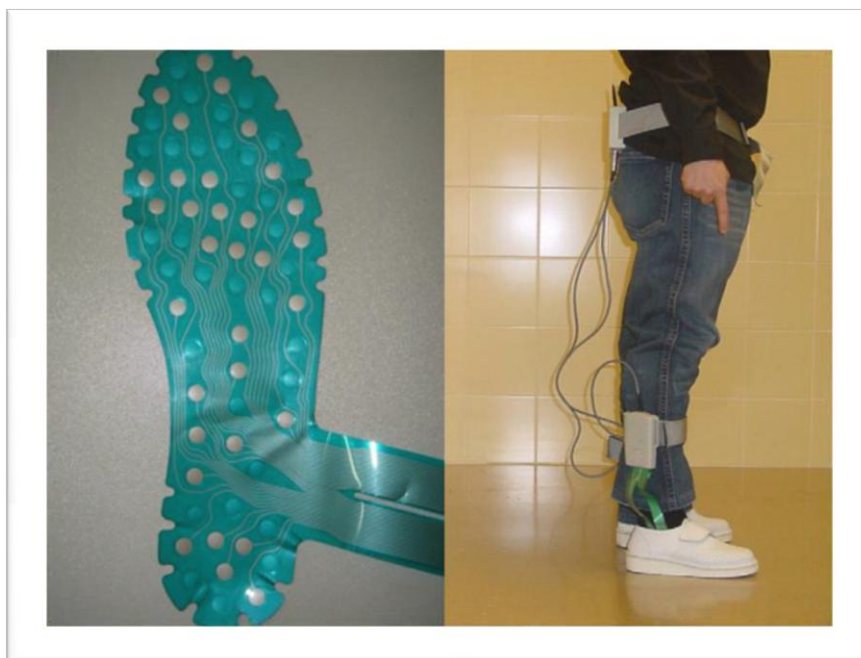
Figura 1 - Transmissão do impulso nervoso entre o cérebro e o pé.  
 Fonte: MERCK, (2011)

Segundo Pedro *et al.* (2015), é possível a criação de soluções que melhorem a vida de pessoas que possuem patologias em que o ajuste na pressão plantar irá melhorar significativamente a qualidade de vida, como, por exemplo, os calçados anatômicos e ortóteses (palmilhas). Em suma, a análise da pressão plantar no pé humano possibilita o diagnóstico de diversas patologias que estejam associadas à postura e aos movimentos do caminhar humano (DEEPASHINI *et al.*, 2014).

### 3.3 SISTEMAS DE MEDIÇÃO

Para a detecção dos dados de pressão plantar, os sistemas de medição mais utilizados e eficazes desenvolvidos são as palmilhas e a plataforma de pressão, que, por suas características são os sistemas com maior aceitação e usabilidade no mercado (HAFER *et al.*, 2013).

As palmilhas (conforme mostrado na Figura 2) são definidas como um objeto flexível com diversos sensores que monitoram as cargas e sua distribuição durante o caminhar, conectado a um amplificador e um sistema de armazenamento (GIACOMOZZI *et al.*, 2012).



**Figura 2 - Sistema de sensores com palmilha.**  
Fonte: Adaptado de MARTINEZ-NOVA *et al.*, (2007).

A palmilha, por sua fácil adaptação, facilita a sua utilização e captação dos dados nas atividades cotidianas, permitindo gravar todo o ciclo de marcha com a mesma máscara. Uma grande vantagem é a análise de pressão em partes do pé que em diversos casos possibilitam a criação de soluções que melhorem a distribuição da pressão plantar (PAPPAS *et al.*, 2014).

Existem algumas desvantagens na utilização das palmilhas, como, principalmente, seu formato e usabilidade, pois seus componentes são bastante visíveis. E no dia a dia é bem notória sua utilização, o tempo de operação das atividades são limitados, dependendo

do sistema de armazenamento, que pode variar de 25 minutos a 6 horas. Além disso, o ambiente fechado no calçado pode comprometer a medição, pois os sensores podem sofrer interferências do suor e calor (SAITO *et al.*, 2011).

A plataforma de pressão (conforme mostrado na Figura 3), diferentemente das palmilhas, são conectadas diretamente ao computador e são capazes de captar toda a distribuição plantar estática e dinâmica por meio dos sensores existentes em sua superfície (LIM *et al.*, 2016).



**Figura 3 - Plataforma de pressões**  
Fonte: SEITZ, (2008).

O registro é feito pelo contato ao caminhar sobre a plataforma, sendo captada a força, sua duração e a quantidade de sensores que foram ativados (LIM *et al.*, 2016). A plataforma, diferentemente da palmilha, consegue medir a força vertical, pois seus sensores estão posicionados paralelamente à superfície; no caso das palmilhas, as pressões do início e do final da captação devem ser excluídas, por serem pressões atípicas que não condizem com a pressão da caminhada normal (BING *et al.*, 2016).

Por possuir maior número de sensores, traz uma visão melhor da diversidade dos resultados. Os fatores externos não influenciam os resultados devido às não interferências mecânicas, e a interação entre o pé e o solo traz uma confiabilidade melhor devido ao próprio formato da plataforma (CASTRO, 2010).

O que pode dificultar a coleta de dados é a marcação da plataforma, que, de certa forma, indica onde serão captados os dados, influenciando o paciente no momento da



pisada e podendo modificar os verdadeiros resultados. Outro problema são os pacientes com patologia severas, que dificultam o seu caminhar e coordenação durante o percurso na plataforma. Essas dificuldades estão realmente relacionadas ao preparo do paciente em executar estas ações, deixando de ser um fato cotidiano de seu caminhar (FERNANDO *et al.*, 2016).

De acordo com Tábuas (2012), as plataformas e as planilhas possuem suas vantagens e desvantagens quanto à sua utilização e obtenção dos dados. Cabe ao profissional analisar que tipo de informação se deseja e as necessidades e impossibilidades de cada paciente para a melhor escolha dos dados de medição da pressão plantar.

O pé humano está condicionado à influência do meio externo, forçando movimentos que possam suprir a necessidade do equilíbrio e movimento do corpo humano. Isso pode comprometer o funcionamento e a forma como são distribuídas as pressões plantares. Esses fatores reforçam a convicção de que tais movimentos podem causar estresse excessivo e o aparecimento de patologias, deformidades plantares ou até mesmo a forma como o corpo exige do pé para conseguir o equilíbrio. Na busca por soluções, a pressão plantar passou ser estudada na investigação de patologias, em que os dados fornecidos possam auxiliar na prevenção, diagnóstico e tratamento de diversas patologias em que o pé esteja associado (TÁBUAS, 2012; PINTO, 2014).

Sistemas de medições foram criados a fim de proporcionar dados úteis na pesquisa dos fatores que influenciam e evidenciam certas patologias. As mais conhecidas e eficazes são as palmilhas e as plataformas de pressão, uma vez que ambas assumiram um papel importante na ciência, tornando os sistemas com maior aplicabilidade no mercado (PINTO, 2014).

Os sistemas de medições como a palmilha e a plataforma de pressão são apenas instrumentos de auxílio que o profissional deve levar em consideração juntamente com as características clínicas e funcionais do paciente para chegar a uma conclusão final do que se pretende observar, cabendo ao profissional a análise dos resultados.

## 4 – DINÂMICA SIMBÓLICA

---

Neste capítulo, são descritas as definições gerais sobre a Dinâmica Simbólica (DS), em seguida descreve-se como são as formas de partição para geração dos símbolos (*Alfabeto*), e finaliza-se o capítulo com a definição das sequências de símbolos (*palavras*) que serão utilizadas para posterior análise simbólica.

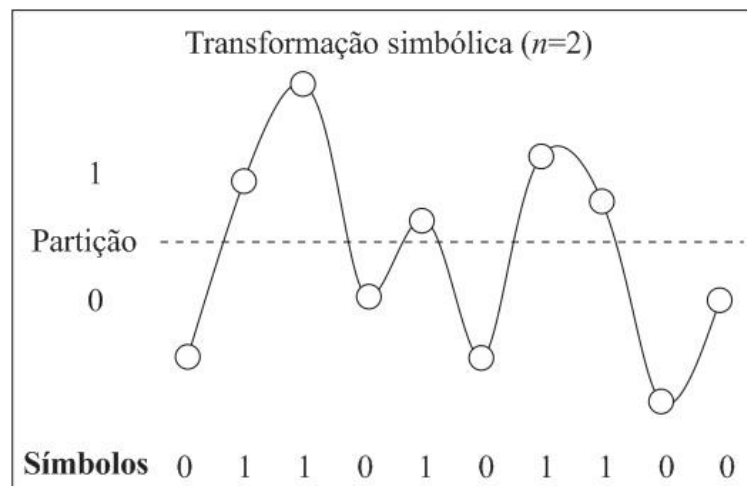
### 4.1 DEFINIÇÕES GERAIS

A realização de medições feitas por meio de experimentos dinâmicos, muitas vezes interessados em testar hipóteses ou fazer inferências com base em padrões temporais nos dados de séries temporais, é um tipo de dinâmica relativamente simples. Para esse tipo de análise de caracterização de padrões, é frequente a aplicação do uso de periodicidade e ferramentas analíticas tradicionais, como a transformada de Fourier (ROCHA, 2013). Mas quando se observa dinâmicas complexas, como bifurcações e oscilações caóticas, esses padrões podem exigir abordagens mais sofisticadas. Nesse caso, o método de análise de dados deve ser selecionado considerando as instalações experimentais e físicas. Detalhes como a resposta dinâmica do instrumento, a taxa de amostragem digital e a relação sinal-ruído podem afetar significativamente a confiabilidade dos resultados.

A Dinâmica Simbólica (DS) é um método que realiza a discretização das medições das séries temporais em uma sequência de símbolos. Esta é tratada como uma transformação dos dados originais, que retêm a maior parte da informação temporal importante. Uma vantagem prática importante do trabalho com símbolos é a facilidade na realização de cálculos numéricos, sendo melhor processados quando comparado aos dados originais. Em alguns casos, essa facilidade pode se manifestar na redução da necessidade de recursos computacionais e aumento da velocidade de resposta por meio do

processamento. Esta última pode ser importante para aplicação de monitoramento em tempo real, além de a análise simbólica de dados ser menos sensível ao ruído (ROCHA, 2013). Em alguns casos, a DS pode ser aplicada diretamente no instrumento de captação utilizando uma combinação adequada de sensores, podendo reduzir significativamente o custo de instrumentação e complexidade da análise (BAUMERT *et al.*, 2015).

A Dinâmica Simbólica é a técnica de modelagem de representação de um sistema dinâmico em função de sequências de símbolos, que representam estados do sistema, e operadores de mudança de estado. Segundo (VALENCIA *et al.*, 2015), a DS descreve a utilização das séries temporais experimentais, na qual um sinal temporal é representado por uma série temporal discreta  $T = \{T_1, T_2 \dots T_n\}$ , em que cada momento no tempo corresponde a um estado representado por um símbolo  $S_i$  que pertence a um alfabeto finito,  $N = \{n_1, n_2, \dots, n_m\}$ . Os estados são, portanto, representados por uma série simbólica  $S = \{S_1 S_2 \dots S_n\}$ . As séries temporais passam por um processo de conversão, no qual os processos de saída acabam assumindo valores discretos e finitos, além de receberem valores utilizando o método de transformação. Por exemplo, utilizando um alfabeto binário ( $n=2$ ), os valores são distribuídos de acordo com um valor que define os símbolos “0” ou “1”, segundo a medida do dado limiar (partição), assim representado pela Figura 4. Seus elementos podem ser agrupados em sequências de  $m$  símbolos consecutivos, totalizando  $n^m$  possíveis combinações de símbolos, ou “palavras” (ROCHA, 2013; SANTOS, 2008; DAW *et al.*, 2003).



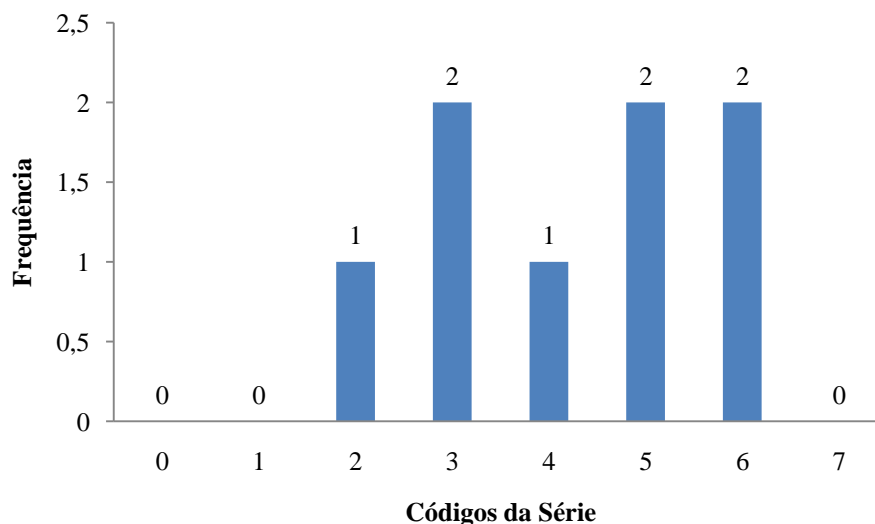
**Figura 4 - Simbolização de série temporal binária (n=2).**

Após simbolizar os dados do sinal, são definidas as sequências de símbolos de tamanho  $m$ . O comprimento  $m$  é representado por um equivalente decimal, por exemplo, uma sequência de  $011_2$  (base binária) passa a ser representada por uma base decimal  $3_{10}$ . Em seguida, são agrupadas as frequências de cada base decimal para a sequência completa obtida da discretização de todas as possíveis sequências de símbolos de tamanho  $m$  (ver Tabela 1).

**Tabela 1 – Sequência de símbolos binários ( $n=2$ ) de comprimento  $m=3$  com seus respectivos códigos de base decimal e frequência de ocorrência no sinal mostrado na Figura 4.**

SEQUÊNCIA DE SÍMBOLOS	CÓDIGO	FREQUÊNCIA DE OCORRÊNCIA
000	0	0
001	1	0
010	2	1
011	3	2
100	4	1
101	5	2
110	6	2
111	7	0

Assim, é possível representar de forma estatística a distribuição de frequência de ocorrência das sequências simbólicas, podendo ser visto por meio de um histograma como na Figura 5.



**Figura 5 - Distribuição das frequências da série temporal da Tabela 1.**

## 4.2 DEFINIÇÃO DE SÍMBOLOS

Para definição dos símbolos que serão utilizados na DS, deve-se particionar o intervalo dos dados originais em um número finito de regiões. Cada região está associada com um valor específico simbólico, e cada medição inicial é mapeada para um determinado símbolo, dependendo da região que a medição represente. O número de símbolos possíveis  $n$  é denominado de *Tamanho do Alfabeto*, de acordo com a literatura simbólica dinâmica. Para o caso mais simples (binário), existem dois símbolos possíveis, onde  $n = 2$ .

Pode-se trabalhar também com diversos valores de  $n$ ; nesse caso, a simbolização terá mais limiares (partições), não sendo representados apenas por “0” e “1”, mas o tamanho ganha diversos outros valores representativos, como o exemplo de  $n=5$ , mostrado na Figura 6.

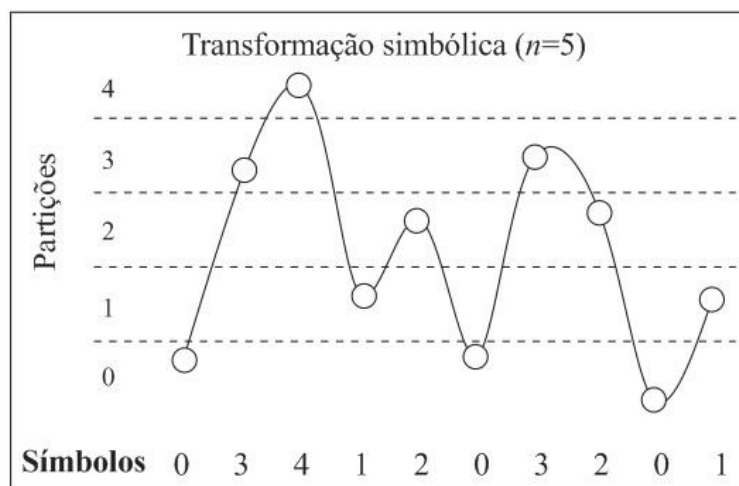


Figura 6 - Transformação simbólica com  $n=5$

Os padrões mais comuns para determinar a localização das partições que irão representar os símbolos são de pontos médios ou medianos, os intervalos de tamanhos iguais ao longo do intervalo de dados ou regiões de intervalo com igual probabilidade (equiprováveis) (DAW & FINNEY, 2002). Por exemplo, Tang *et al.* (1995 *apud* SANTOS, 2008) e Tang *et al.* (1997 *apud* SANTOS, 2008) verificaram que um conjunto de símbolos binários divididos sobre uma amostra média foi adequado para reconstruir a dinâmica dos modelos não lineares, mesmo quando a dinâmica observada estava fortemente contaminadas com ruído. Rapp *et al.* (1994 *apud* KABIR *et al.*, 2011)

investigaram erros na utilização de particionamento por ponto médio, em vez de particionamento pela média dos valores, e sugeriu uma maneira para verificar se há identificações falsas de estruturas não aleatórias. Hively *et al.* (2000 *apud* DAW *et al.*, 2002), por outro lado, utilizaram intervalos de tamanhos iguais para dividir os sinais de eletroencefalograma a fim de detectar precursores para convulsões. KIM *et al.* (2000 *apud* SANTOS, 2008) analisaram a dinâmica de frequência cardíaca usando partições alinhadas com os dados médios e desvios padrões das amostras. Santos (2008) realizou análise estatística das sequências de símbolos de um sistema dinâmico para diferenciar dinâmicas determinísticas e comportamentos estocásticos. Para isso foi realizada uma partição do espaço de estados do sistema denominada partição estatística. Esse tipo de partição é geralmente utilizada quando o sinal analisado é contaminado por ruído, impedindo a definição de uma partição topológica (ROCHA, 2013).

No entanto, partições são selecionadas de acordo com a sensibilidade dos resultados, indicando que a escolha da partição deve ser cuidadosamente avaliada. É possível também escolher locais de partições *danificadas*, onde a maior parte da informação dinâmica significativa é perdida (SOUZA, 2012; VALENCIA *et al.*, 2015).

### 4.3 CODIFICAÇÃO

Após o particionamento e posteriormente a simbolização dos dados, o passo seguinte na identificação de padrões temporais é a construção da sequência de símbolos ou *Palavras* (de acordo a literatura da Dinâmica Simbólica). Esta série é o agrupamento de símbolos em ordem temporal. Esse processo sequencial é a definição de um modelo de comprimento finito, que pode ser movido ao longo da série de símbolos um passo de cada vez, revelando uma nova sequência a cada etapa. Cada possível sequência é representada por um identificador único, em que o resultado final será uma nova série temporal, referida como uma sequência de série de símbolos ou código da série (ROCHA, 2013; DAW *et al.*, 2003).

A Figura 7 ilustra o processo de transformação simbólica de uma série temporal convertida em série simbólica binária ( $n=2$ ). No exemplo, as sequências de símbolos são agrupadas sucessivamente de três em três símbolos binários ( $m=3$ ). Cada sequência possível é representada por seu valor binário ou valor decimal correspondente, determinado pela posição de cada símbolo no modelo.

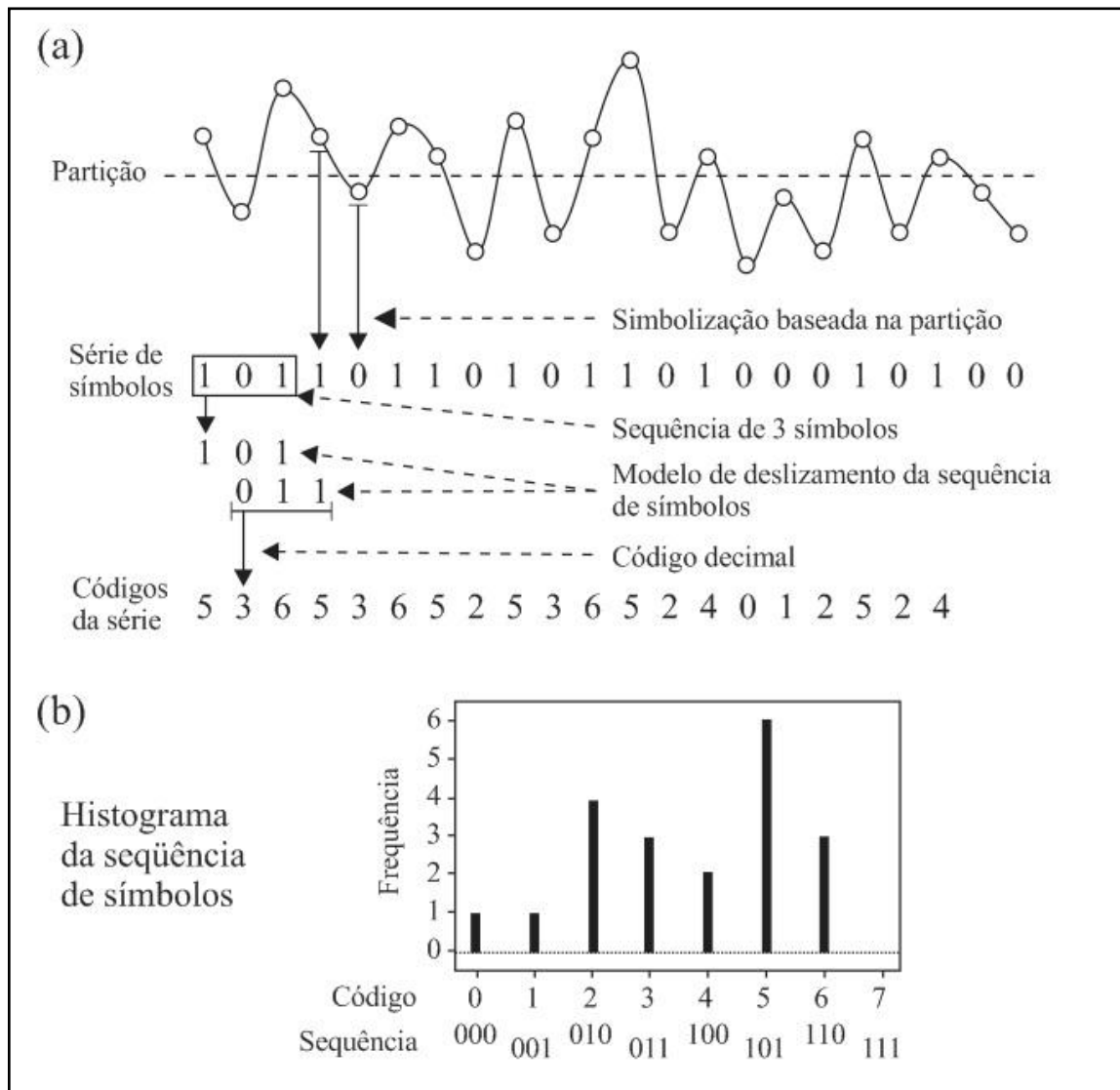


Figura 7 – (a) Processo de simbolização de série temporal; (b) Tabulação de um histograma da sequência simbólica

Fonte: Adaptado de DAW & FINNEY, (2002).

Além disso, é possível realizar análises estatísticas das sequências de símbolos geradas, de forma a se estimar a probabilidade de ocorrência de diferentes sequências (palavras) de tamanhos limitados. Assim, por meio de análises, é possível encontrar sequências limitadas que caracterizem uma determinada dinâmica ou diferenciem os tipos distintos de dinâmicas presentes em um sistema em diferentes condições (KABIR *et al.*, 2011).

# 5 – ALGORITMO GENÉTICO

---

Este capítulo é iniciado com uma breve introdução sobre os Algoritmos Genéticos (AGs), em seguida destacam-se os elementos que os compõem e descrevem-se os principais operadores genéticos e a definição dos parâmetros genéticos.

## 5.1 INTRODUÇÃO AO ALGORITMO GENÉTICO

A otimização é um paradigma fundamental em muitas aplicações, tais como engenharia, atividades empresariais, na indústria, tomada de decisão de preço de produtos de forma a maximizar o lucro, ou a escolha de um trajeto de ônibus que maximize o número de passageiros servidos e minimize as distâncias percorridas, entre outras. Naturalmente, os objetivos de otimização, em todas as aplicações citadas acima, são a minimização do consumo de energia e custos, ou a maximização do lucro, da produção, do desempenho e a eficiência. Na realidade, os recursos (tempo e dinheiro) são sempre limitados e, conseqüentemente, a otimização destes é um fator essencial (ANOCHI, 2015).

A própria evolução das espécies pode ser vista como um processo de otimização: ao longo do tempo, os seres vivos se tornam cada vez mais bem adaptados a um meio ambiente em constante mudança.

Considerando o espaço de busca a ser explorado, os problemas de otimização são tratados por diversas áreas do conhecimento. Dentre elas são citadas (BECCENERI, 2013):

- Programação linear: a função custo (objetivo) é linear e há um conjunto de restrições relacionadas às variáveis da função, que são especificadas como igualdades ou desigualdades lineares;



- Otimização contínua: as variáveis assumem valores reais ou contínuos;
- Otimização combinatória (discreta): as variáveis assumem valores discretos ou inteiros;
- Otimização mista: com variáveis inteiras e contínuas ao mesmo tempo;
- Otimização estocástica: estuda os casos em que as restrições ou os parâmetros dependem de variáveis aleatórias.

Candan & Yazgan. (2014) apresentam alguns conceitos básicos de otimização:

- Variáveis de decisão: formam um conjunto de  $n$  parâmetros cujos valores representam uma solução para um problema de otimização. Elas podem ser contínuas (reais), inteiras ou discretas (valores compreendidos dentro de um conjunto fixo).
- Restrições: são funções das variáveis de decisão que delimitam o problema e validam as soluções. Pode-se dizer que as restrições esboçam o contorno da região onde se encontra o conjunto de soluções factíveis do problema.
- As restrições dividem-se em dois grupos: restrições laterais onde são efetuadas diretamente sobre as variáveis de decisão, limitando seus valores; e a restrição de comportamento que consiste nas condições desejáveis de limites de tensões, deslocamentos, frequências naturais de vibração, entre outros.
- Funções objetivo: são critérios de avaliação para determinar a qualidade de uma solução, ou seja, é a função que se deseja otimizar. Na otimização multiobjetivo existem duas ou mais funções objetivo em cada problema.
- Ponto Ótimo: é o ponto caracterizado pelo vetor  $x = (x_1, x_2, \dots, x_n)$  formado pelas variáveis de decisão que otimizam as funções objetivo e satisfazem as restrições.
- Valor Ótimo: é o valor da função objetivo  $f(x)$  no ponto ótimo.
- Solução Ótima: a solução ótima para cada função  $j$  é o par formado pelo ponto ótimo e o valor ótimo  $[x^*, f(x^*)]$ .

Os métodos de otimização baseados em derivadas necessitam que as funções sejam contínuas e diferenciáveis, podendo ficar presos a mínimos locais, fato esse que limita seu domínio de aplicação. Os métodos estocásticos, por sua vez, podem trabalhar tanto com codificação contínua como discreta das variáveis e não necessitam que as funções sejam diferenciáveis (SOUZA, 2009). Entretanto, nem sempre é possível obter uma expressão analítica para a derivada da função objetivo, ou o cálculo do gradiente pode ser muito oneroso sob o ponto de vista computacional. Nesse caso, os métodos utilizando Algoritmos Evolucionários se apresentam como uma ferramenta útil, estes algoritmos exigem muito pouco da função objetivo (em relação à sua estrutura) e prosperam em ambientes multimodais e até mesmo descontínuos. Além disso, com a codificação adequada, eles podem manipular variáveis inteiras e contínuas ao mesmo tempo e ainda lidar com restrições lineares e não lineares trivialmente (PAULA, 2015).

Dentro do domínio dos procedimentos de busca e otimização encontra-se o Algoritmo Genético (AG). Os AGs são inspirados no processo de evolução natural e utilizados para resolver problemas reais de otimização. Segundo Malaquias (2006), eles procuram a melhor solução para os problemas, utilizando um processo iterativo de busca.

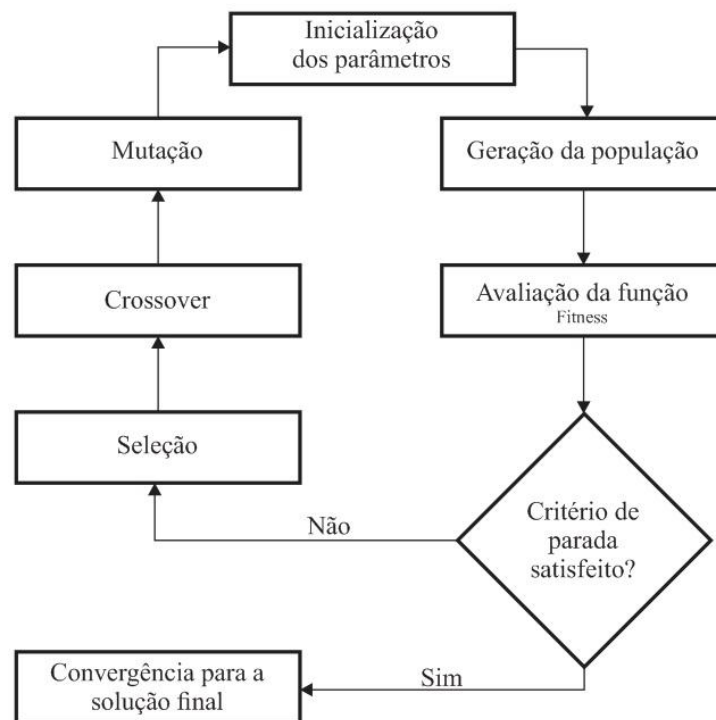
Diferente dos métodos tradicionais de procura da solução ótima, os AGs apresentam algumas peculiaridades que lhes conferem a possibilidade de explorar espaços de procura mais complexos de forma robusta, como trabalharem com uma codificação das possíveis soluções (genótipos) e não com as próprias soluções (fenótipos), manipularem simultaneamente um conjunto (população) de soluções candidatas, utilizarem informações de custo e recompensa e, por fim, empregarem regras de transição probabilísticas (SOUZA, 2015; CANDAN & YAZGAN, 2014).

Os AGs permitem uma simplificação na formulação e solução de problemas de otimização, pois incorporam uma solução potencial para um problema específico numa estrutura semelhante à de um cromossomo, e aplicam operadores de seleção e cruzamento a essas estruturas de forma a preservar informações críticas relativas à solução do problema (ANOCHI, 2015).

Normalmente os algoritmos genéticos são vistos como otimizadores de funções, embora a quantidade de problemas para a qual os algoritmos genéticos se aplica seja bastante abrangente (GOMES, 2013). A Programação Evolutiva é um algoritmo genético

probabilístico iterativo que mantém, a cada iteração  $t$ , uma população de indivíduos  $P(t)$ :  $P(t) = (p_1^t, p_2^t, \dots, p_n^t)$ . Em termos matemáticos, cada indivíduo  $p^t$  representa uma solução do problema associado. A cada iteração  $t$  existe um mecanismo que permite a renovação da população, obtendo  $P(t)$  a partir de  $P(t - 1)$ . Também a cada iteração  $t$  cada indivíduo  $p^t$  é avaliado segundo uma função que mede o nível da sua aptidão por critérios pré-definidos. Dessa forma, aqueles indivíduos considerados mais aptos sobrevivem, ou seja, passam para a população da iteração seguinte, e aqueles considerados menos aptos, serão descartados (ALAJMI & WRIGHT, 2014).

Esses procedimentos conduzem a um processo de renovação iterativa das populações de modo a tentar melhorar as qualidades genéticas de cada indivíduo. Um programa em AG segue o fluxo ilustrado pela Figura 8.



**Figura 8 - Fluxograma de otimização utilizando o Algoritmo Genético.**  
**Fonte: Adaptado de PAULA, 2015.**

Preliminarmente, é escolhida uma população inicial e a qualidade dessa população é determinada. Em seguida, em cada iteração, pais são escolhidos da população para produzir filhos, que são adicionados à população. Cada indivíduo da população resultante pode então sofrer alguma mutação, que é uma alteração aleatória no cromossomo (COSTA, 2015).

## 5.2 ELEMENTOS DO ALGORITMO GENÉTICO

Antes de explorarmos os elementos existentes nos algoritmos genéticos, é preciso ressaltar que não há uma definição única e rigorosa aceita por todos que fazem uso da computação evolutiva, diferenciando esta técnica das demais. Porém, alguns elementos costumam ser padrões nos AGs – Dentre eles destacam-se a população representada por cromossomos, a seleção de indivíduos dependentes de sua aptidão, a função objetivo para o cálculo da aptidão, a recombinação e mutação para produzir uma nova geração e o aumento da variabilidade genética na população – (GOMES, 2013; LIM *et al.*, 2015).

O procedimento de um AG começa pela geração de uma população inicial de indivíduos (cromossomos). Durante a execução do AG, essa população é avaliada, de modo que cada indivíduo receba uma nota (chamada de aptidão do indivíduo), proporcional à qualidade da solução a que ele corresponde. Normalmente os indivíduos mais aptos são selecionados, podendo ser alterados por meio de recombinação e mutação durante a reprodução, produzindo descendentes para a próxima geração. Já os indivíduos menos aptos da população costumam ser descartados. Esse mecanismo reflete a seleção natural proposta por Darwin e é repetido até que algum critério de convergência, como tempo de simulação, número de gerações ou valor de aptidão seja atingido (COSTA, 2015). A Figura 9 exibe um AG típico utilizado para representar genericamente sua aplicação em um ambiente de programação computacional.

```
Seja  $S(t)$  a população de cromossomos na geração  $t$  (comentário)
 $t \leftarrow 0$ 
inicializar  $S(t)$ 
avaliar  $S(t)$ 
enquanto o critério de parada não for satisfeito faça
     $t \leftarrow t + 1$ 
    selecionar  $S(t)$  a partir de  $S(t - 1)$ 
    aplicar recombinação sobre  $S(t)$ 
    aplicar mutação sobre  $S(t)$ 
    avaliar  $S(t)$ 
fim enquanto
```

**Figura 9 - Algoritmo Genético típico**  
**Fonte: Adaptado de COSTA, (2015).**

Cada um dos elementos presentes no AG será explicado a seguir.

### **5.3 REPRESENTAÇÃO E CODIFICAÇÃO**

Basicamente, existem três tipos de codificação dos parâmetros em algoritmos genéticos. A codificação clássica (ou binária), números inteiros e a codificação real (SOUZA, 2009). Esta codificação é necessária para transformar as variáveis do problema em um cromossomo para que o algoritmo possa manipular corretamente.

A codificação binária, ou clássica, foi a primeira a ser construída, sendo utilizada na maioria dos trabalhos desenvolvidos (HOSSEINI *et al.*, 2014; GOPALAKRISHNAN & KOSANOVIC, 2015; SANTOS *et al.*, 2014). Na codificação binária, cada cromossomo é um vetor composto por zeros e uns, com cada bit representando um gene do mesmo. Porém, conforme a necessidade de maior precisão numérica, as cadeias de bits tornam-se excessivamente longas, o que acarreta a necessidade de um esforço computacional maior, causando também um consumo maior de tempo até a convergência dos algoritmos (LIM *et al.*, 2015). A codificação real trabalha diretamente com os números reais, o que torna possível cobrir um domínio bastante abrangente, mesmo para domínios desconhecidos, das variáveis. Utilizam-se números de ponto flutuante para representar o valor das variáveis e executar as operações genéticas de cruzamento e mutação. Existe, assim, uma grande desvantagem quanto a representação binária, que seriam as constantes conversões entre valores de ponto flutuantes (PAULA, 2015).

### **5.4 POPULAÇÃO**

O tamanho da população inicial e das populações subsequentes é um aspecto importante a ser considerado para o sucesso do processo evolutivo, pois é possível comprometer a evolução dos indivíduos caso a quantidade escolhida seja muito grande ou muito pequena. Essa medida, exceto em casos muito particulares, deve ser estabelecida empiricamente e de acordo com a disponibilidade de recursos computacionais (GOMES, 2013).

A população normalmente é inicializada aleatoriamente, porém processos heurísticos também podem ser utilizados, além da utilização na primeira população dos valores resultantes de outros métodos de otimização, sendo esta última uma técnica

vantajosa de destaque para os AGs, pois inicia o processo de otimização com soluções obtidas por outros métodos (COSTA, 2015).

## 5.5 SELEÇÃO

Inspirado no processo de seleção natural de seres vivos, o algoritmo genético seleciona indivíduos da população, geralmente aqueles com alta aptidão para gerar cromossomos filhos (variantes dos pais) por meio dos operadores de recombinação e mutação. Os descendentes constituem a geração seguinte. Geralmente, os pais são selecionados com probabilidade proporcional à sua aptidão para gerar descendentes. Existem muitos métodos para fazer essa seleção. Exemplos são: seleção por roleta, por torneio, por elitismo, por classificação, por estado estacionário, entre outros. Descreveremos os métodos de seleção mais comumente utilizados, que são: seleção por elitismo, por roleta e por torneio (GOPALAKRISHNAN *et al.*, 2015).

**Seleção por elitismo:** neste tipo de seleção,  $k$  indivíduos com as melhores aptidões da população são selecionados para continuarem na próxima geração sem serem alterados, ou seja, sem passarem pelos operadores de reprodução (recombinação e mutação). Isso evita que possíveis boas soluções sejam perdidas. É comum utilizar  $k = 1$ ; porém, aumentando o valor de  $k$ , aumenta-se a pressão seletiva, isto é, o risco de o sistema convergir prematuramente a uma solução que não seja necessariamente a solução ótima para o problema.

**Seleção por roleta:** neste método, cada indivíduo da população é representado em uma roleta proporcionalmente à sua aptidão. Assim, aos indivíduos com alta aptidão é dada uma porção maior da roleta, enquanto aos de aptidão mais baixa é dada uma porção relativamente menor da roleta. Finalmente, a roleta é girada um determinado número de vezes escolhido, dependendo do tamanho da população, e aqueles indivíduos nela sorteados são selecionados para se reproduzirem, deixando descendentes para a próxima geração.

Na prática, a aptidão de todos os indivíduos da população é somada e normalizada. Um valor aleatório entre zero e um é sorteado e é a aptidão relativa de cada indivíduo que determina a faixa de valores que representa cada um deles, como ilustra o exemplo contido na Figura 10. Esse processo pode ser realizado de maneira equivalente, sorteando um

número aleatório entre zero e a soma da aptidão de todos os indivíduos; nesse caso, a faixa de valores que representa cada membro da população depende diretamente da aptidão deles e não mais da aptidão relativa. Na seleção por roleta, a pressão seletiva é menor do que na seleção por elitismo.

Outro método de seleção é semelhante à roleta, chamado de seleção por classificação (ou por *ranking*), com a diferença de que a probabilidade de seleção é relacionada também à ordenação dos indivíduos conforme sua aptidão antes do sorteio. A ordenação pode ser, por exemplo, decrescente, linear ou exponencial.

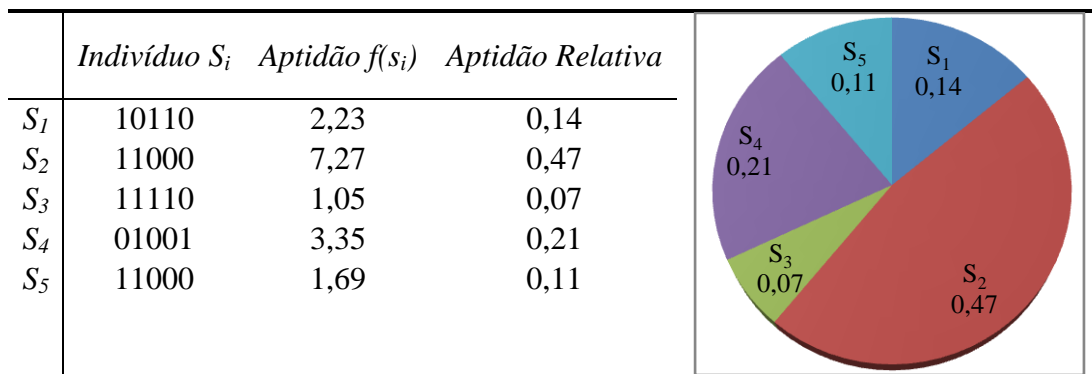


Figura 10 - Indivíduos de uma população e a sua correspondente roleta de seleção  
 Fonte: Adaptado de COSTA, (2015).

**Seleção por torneio:** existem dois tipos de implementação do torneio. No primeiro, dois indivíduos são sorteados e o melhor deles é escolhido com probabilidade  $p_s$ . No segundo,  $k$  indivíduos são sorteados e o melhor deles é escolhido. Aumentando-se  $p_s$  ou diminuindo-se  $k$ , aumenta-se a pressão seletiva. A possibilidade de se controlar a pressão seletiva é a principal vantagem do método. O processo de torneio é repetido até que o número de pais que originarão a população da geração seguinte seja atingido.

É possível, e é o que normalmente se faz, combinar mais de um tipo de seleção para gerar a próxima população, aproveitando as vantagens e suavizando as desvantagens de cada método (COSTA, 2015). A etapa de seleção é responsável por determinar quais indivíduos serão escolhidos para a etapa de reprodução. A probabilidade de uma dada solução ser escolhida para a próxima etapa é proporcional ao grau de aptidão dessa solução. A seleção se baseia no princípio da “sobrevivência do mais apto”, em que as soluções boas são selecionadas para a reprodução e as soluções ruins são eliminadas (COSTA, 2015; CANDAN & YAZGAN, 2014).

## 5.6 FUNÇÃO DE APTIDÃO

O propósito da otimização realizada pelo AG é maximizar (ou minimizar) uma função de aptidão (também chamada de adequação, custo, objetivo ou *fitness*). Sendo assim, um dos aspectos mais importantes do desenvolvimento de AGs é a escolha de uma função de aptidão apropriada (HOLLAND, 1975). A população é avaliada durante o processo de evolução do algoritmo. É conferida uma classificação para cada indivíduo, que reflete o grau de adaptação deste em relação ao ambiente (SOUSA, 2015). Segundo Li *et al.* (2014) este tipo de algoritmo requer um valor escalar que represente o desempenho de cada solução. Essa avaliação é então baseada em uma função de aptidão, conhecida por *fitness*, que é obtida por meio de uma função a ser otimizada ao avaliar o cromossomo (indivíduo), geralmente a função objetivo do problema de otimização. Logo, a aptidão é um valor obtido com a aplicação desta função, que pode ser conhecido pela quantificação da aptidão do indivíduo (SOUSA, 2015). Segundo Holland (1975), a função *fitness* é uma indicação de quão desejável é um indivíduo, em termos de sobrevivência para a próxima geração.

## 5.7 CRUZAMENTO

O cruzamento (também denominado na literatura como *Recombinação* e *Crossover*) é o elemento que distingue os AGs das outras estratégias em computação evolutiva (COSTA, 2015). É o operador responsável pela recombinação de características dos pais durante a reprodução, permitindo que as próximas gerações herdem essas características. A ideia principal é propagar as características positivas dos indivíduos mais aptos da população por meio da troca de segmentos de informações entre os mesmos, o que originará novos indivíduos. As formas mais comuns de troca de segmentos nos AG são (ANOCHI, 2015):

**Um - ponto:** O ponto de cruzamento é escolhido de forma aleatória e a partir deste ponto as informações genéticas dos pais serão trocadas. As informações anteriores a esse ponto em um dos pais são ligadas às informações posteriores a ele no outro pai, como é mostrado na Figura 11.



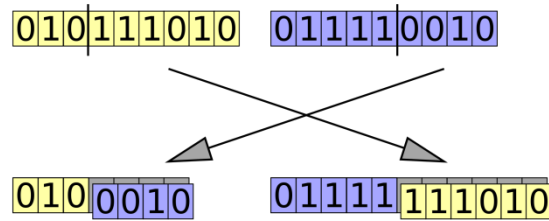


Figura 11 - Operador de cruzamento em um-ponto

**Multipontos:** é realizada da mesma forma do cruzamento de um ponto, porém a troca de material genético através de um número maior de pontos, onde muitos pontos de cruzamento podem ser utilizados, conforme ilustrado na Figura 12.

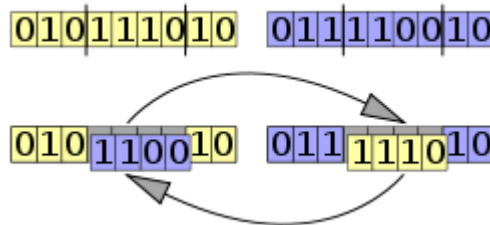


Figura 12 - Operador de cruzamento multiponto

## 5.8 MUTAÇÃO

A mutação é o fator fundamental para garantir a biodiversidade e introduzir aleatoriedade na população. Apesar de normalmente aplicada com uma probabilidade bastante inferior à de cruzamento, ela é tida por uma série de autores como o operador mais importante para o processo evolutivo, chegando, em alguns casos extremos, a ser utilizada como o único operador de transformação (SOUSA, 2015; PIRES, 2010).

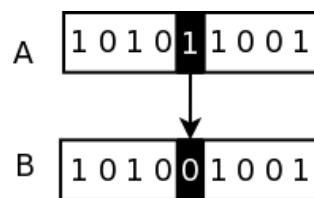
A operação de mutação é responsável por aplicar pequenas mudanças aleatórias a um ou vários genes de um cromossomo, a fim de promover variação e diversidade na população. A mutação amplia o espaço de busca, o que aumenta a exploração da superfície de resposta (GOMES, 2013).

O operador de mutação possui também um papel fundamental no que diz respeito à necessidade de evitar a convergência prematura, que ocorre quando a população se estabiliza com uma média de adaptação pouco adequada por causa da pressão evolutiva e baixa diversidade. Isso geralmente dá-se com o surgimento de um superindivíduo que

domina o processo seletivo e, uma vez incapaz de gerar filhos melhores, transmite as suas características por toda a população.

Segundo Alajmi & WRIGHT (2014) o papel da mutação é prover a oportunidade de recuperação de alguma característica genética boa que possa ter sido perdida durante as etapas de seleção e cruzamento. Candan & YAZGAN (2014) afirmam que, apesar de o operador ser considerado um operador secundário por alguns autores, a operação de mutação é a única operação que introduz informação genética não representada na população inicial, fator essencial para que o algoritmo não fique preso a um ótimo local.

O procedimento de mutação na codificação binária consiste na alteração de um bit de 0 para 1 (e vice-versa), como exemplificado na Figura 13.



**Figura 13 - Operador de mutação**  
**Fonte: Adaptado de SOUZA, 2009.**

Existem diferentes tipos de mutação que são responsáveis por determinar quais bits sofrerão mutação; isso ocorre quando a representação é feita por meio de números inteiros ou reais, e para isso métodos mais complexos são necessários para modificar um gene. Nesses casos, é comum se usar as seguintes estratégias (COSTA, 2015; SANTOS *et al.*, 2014):

- **Mutação uniforme:** substitui o gene por um número aleatório dentro dos limites permitidos para ele;
- **Mutação de limite:** substitui o gene por um dos limites do intervalo factível a ele;
- **Mutação não-uniforme:** substitui um gene por um número extraído de uma distribuição não uniforme (seja linear, exponencial, Lorentziana, etc.);
- **Mutação gaussiana:** é um método bastante utilizado. Substitui o gene por um número aleatório de uma distribuição gaussiana com média e desvio padrão determinados no AG;
- **Mutação creep:** adiciona ao parâmetro um pequeno valor aleatório obtido a partir de distribuição uniforme ou normal.

## 5.9 PARÂMETROS GENÉTICOS

Os parâmetros genéticos influenciam o comportamento e desempenho dos algoritmos genéticos, e afetam da mesma forma a codificação real e a codificação clássica (SOUZA, 2009). Após a apresentação de todos os elementos do AG, vale dizer que não há valores ideais para cada parâmetro do modelo que garantam bons resultados para todos os tipos de problemas (COSTA, 2015). O mais importante ao implementar o AG é equilibrar todos os parâmetros entre si, conforme o problema que se pretende solucionar (PAULA, 2015). A eficiência dos Algoritmos Genéticos é altamente dependente dos seguintes parâmetros:

**Tamanho da população:** o tamanho da população afeta o desempenho global e a eficiência dos AG. Ao trabalhar-se com uma população pequena, o desempenho pode ser comprometido, pois desse modo o espaço de busca é reduzido. Uma grande população geralmente fornece uma cobertura representativa do domínio do problema, além de prevenir convergências prematuras para soluções locais ao invés de globais. Porém, ao trabalhar com grandes populações, são necessários maiores recursos computacionais (SOUZA, 2009).

**Taxa de cruzamento:** uma alta taxa de cruzamento (ou *crossover*) permite que novas estruturas sejam reveladas na população. No entanto, se essa taxa for muito elevada, a maior parte da população com boas aptidões serão substituídas (SOUZA, 2009). Um valor baixo, o algoritmo pode tornar o algoritmo muito lento (COSTA, 2015).

A taxa de cruzamento se refere à porcentagem da população de pais que irão sofrer a operação de cruzamento. Segundo Paula (2015), enquanto uma taxa de cruzamento alta pode causar o descarte de boas soluções, uma taxa muito baixa pode dar muita atenção aos cromossomos pais e então estagnar a busca. De acordo com Candan *et al.* (2014), quanto maior a taxa de cruzamento, mais rapidamente novas estruturas são reveladas na população.

**Taxa de mutação:** a taxa de mutação se refere à porcentagem da população de pais que irão sofrer a operação de mutação.

Segundo Paula (2015), Souza (2009) e Costa (2015), pode-se observar que para uma taxa de mutação muito baixa, pode acontecer que a busca fique estagnada em um

ótimo local. Em contrapartida, com uma taxa muito alta, a busca torna-se essencialmente aleatória.

**Número de gerações:** finalmente é necessário decidir quando parar o algoritmo. Existem dois critérios para esta tomada de decisão (SOUSA, 2015):

- Parar após um número pré-definido de gerações ou um limite de tempo pré-definido;
- Parar quando nenhuma melhoria for mais conseguida depois de um determinado número de gerações, ou seja, quando o valor do melhor *fitness* em duas gerações seguidas, por exemplo, for inferior ao melhor valor já atingido.

## 5.10 APLICAÇÃO

Algoritmos genéticos são complexos e nem sempre são o método mais eficiente para solucionar um problema. Embora existam muitos exemplos de sucesso do uso de AGs, em outros casos eles funcionam pobremente. Não existe uma resposta única sobre quando convém serem utilizados, mas muitos pesquisadores discutem que podem ser comparáveis ou melhores que outras técnicas de otimização, principalmente nas situações em que o espaço de busca seja multimodal, que a função objetivo seja ruidosa e que o encontro de um ótimo local da função seja suficiente, não carecendo necessariamente do ótimo global (COSTA, 2015). Isso porque um espaço de busca grande e desconhecido inviabiliza sua varredura exaustiva, pois a melhor solução não pode ser rapidamente encontrada. As principais vantagens dos algoritmos genéticos são, portanto: robustez, pouca influência em seu desempenho decorrente de descontinuidades no espaço de busca e não requisição de conhecimentos prévios dos gradientes da função objetivo (PAULA, 2015; LIM *et al.*, 2015; GOPALAKRISHNAN *et al.*, 2015).

Assim, selecionamos os algoritmos genéticos para otimizar os parâmetros da Dinâmica Simbólica, dadas as qualidades supramencionadas para esse método.

# 6 – REDES NEURAIIS ARTIFICIAIS PARA CLASSIFICAÇÃO

---

Neste capítulo é apresentada uma breve introdução às Redes Neurais Artificiais. São discutidos os conceitos básicos das arquiteturas das RNAs, assim como os algoritmos de aprendizado, as fases de treinamento e os tipos de redes neurais, com ênfase na descrição das redes do tipo Perceptron de Múltiplas Camadas (MLP) utilizada neste trabalho.

## 6.1 INTRODUÇÃO ÀS REDES NEURAIIS ARTIFICIAIS

Redes Neurais Artificiais (RNAs) são sistemas paralelos e distribuídos compostos por unidades de processamento simples (nodos ou neurônios ou ainda, unidades) que calculam determinadas funções matemáticas (normalmente não lineares). Tais unidades são dispostas em uma ou mais camadas e interligadas por um grande número de conexões, geralmente unidirecionais. Na maioria dos modelos, essas conexões estão associadas a pesos, os quais armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio da rede (HASSAN *et al.*, 2016).

Segundo Kristjanpoller & Minutolo (2015), a solução de problemas por meio de RNAs é bastante atrativa, já que a forma como estes são representados internamente pela rede e o paralelismo natural inerente à arquitetura das RNAs criam a possibilidade de um desempenho superior ao dos modelos convencionais. Em RNAs, o procedimento usual na solução de problemas passa inicialmente por uma fase de aprendizagem, em que um conjunto de exemplos é apresentado para a rede, a qual extrai automaticamente as

características necessárias para representar a informação fornecida. Essas características são utilizadas posteriormente para gerar respostas para o problema (FLORES, 2009).

A capacidade de aprender por meio de exemplos e de generalizar a informação aprendida é o atrativo principal da solução de problemas por RNAs. A generalização, que está associada à capacidade da rede aprender por intermédio de um conjunto reduzido de exemplos e posteriormente dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das RNAs vai muito além do que simplesmente mapear relações de entrada e saída. As RNAs são capazes de extrair informações não apresentadas de forma explícita por meio dos exemplos (MORAIS, 2010).

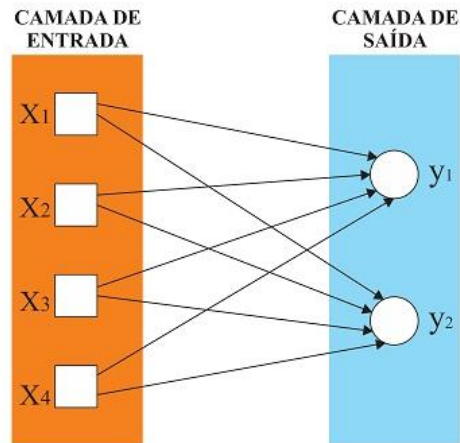
As RNAs tentam reproduzir as funções das redes biológicas, buscando implementar seu comportamento básico e sua dinâmica. No entanto, do ponto de vista físico, no momento as redes artificiais se diferenciam bastante das redes biológicas. É importante, contudo, observar as similaridades entre esses dois tipos de sistemas, tanto para que se possa entender melhor o sistema nervoso quanto para buscar idéias e inspirações para a pesquisa em neurocomputação (MORAIS, 2010). Como características comuns, pode-se citar que os dois sistemas são baseados em unidades de computação paralela e distribuída, que se comunicam por meio de conexões sinápticas, possuem detectores de características, redundância e modularização das conexões (MENEZES *et al.*, 2015).

## 6.2 ARQUITETURA DE RNA

Em uma RNA, os neurônios são organizados em camadas, sendo a arquitetura da rede definida pelo número de camadas, disposição dos neurônios em cada camada, topologia das conexões e tipo de função de ativação dos neurônios. O número mínimo de camadas que uma RNA pode possuir é dois: uma camada de entrada e uma de saída (COSTA, 2015). Os modelos de RNA destacam-se quanto ao formato e à topologia de rede, sendo dividido em três classes gerais, diferentes no seu funcionamento (HAYKIN, 2001; ANOCHI, 2015; LARRAIN & TURNER, 2014): (i) redes de camada única, (ii) redes com múltiplas camadas e (iii) redes recorrentes.

**Redes de camada única:** é considerado o mais simples modelo de rede (HAYKIN, 2001; SOARES *et al.*, 2014; MARTINIANO *et al.*, 2016). Redes com essa arquitetura têm apenas a camada de entrada e a camada de saída, não possuindo nenhuma camada oculta. É

alimentada adiante, pois o fluxo de informação é unidirecional, sendo projetado apenas da camada de entrada para a camada de saída. A Figura 14 apresenta um exemplo de rede alimentada adiante com uma camada única.



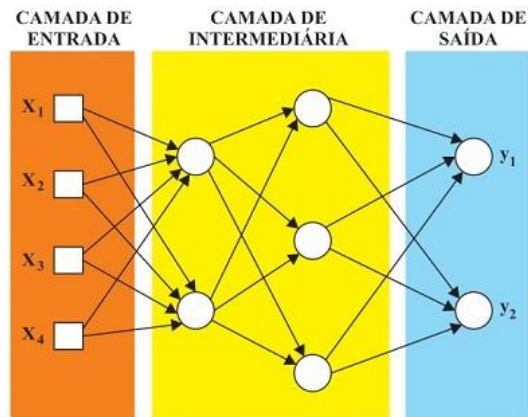
**Figura 14 - Arquitetura de rede de camada única**  
Fonte: Adaptado de FLORES, 2009.

**Redes com múltiplas camadas:** A rede *Perceptron* de Múltiplas Camadas surgiu como alternativa para a solução de problemas não linearmente separáveis e tem sido aplicada com sucesso para solução de diversos problemas complexos, por meio do seu treinamento supervisionado, usando o algoritmo de retropropagação de erro (*error backpropagation*) baseado na regra de aprendizagem por correção do erro (RUSSEL & NORVIG, 2013).

Os sinais de entrada são propagados camada por camada até que o vetor de saída seja obtido na última camada. O processo de aprendizado termina quando o erro desejado é atingido ou quando se atinge um número máximo de épocas de aprendizado (RATH *et al.*, 2015). A grande vantagem de redes de camadas múltiplas é a complexidade dos sistemas com que elas são capazes de lidar. Isso se deve ao aumento de capacidade de sinapse oferecido pelas camadas ocultas e camada de saída. Como o sinal de cada camada oculta depende da camada imediatamente anterior, isso aumenta a capacidade da rede. Isso também aumenta a quantidade de parâmetros do modelo (COSTA, 2015; ARRUDA *et al.*, 2013).

A arquitetura de uma rede de múltiplas camadas consiste no arranjo topológico das unidades de processamento dos neurônios com os respectivos valores de pesos associados às conexões. Os pesos sinápticos são ajustados para fazer com que a resposta real da rede

se aproxime da resposta desejada (MARTINIANO *et al.*, 2016). A Figura 15 mostra uma RNA multicamada com duas camadas intermediárias. Os círculos indicam neurônios, os quais processam sinais vindos de suas entradas, segundo uma arquitetura que pode variar de modelo para modelo. Os quadrados indicam as unidades sensoriais.

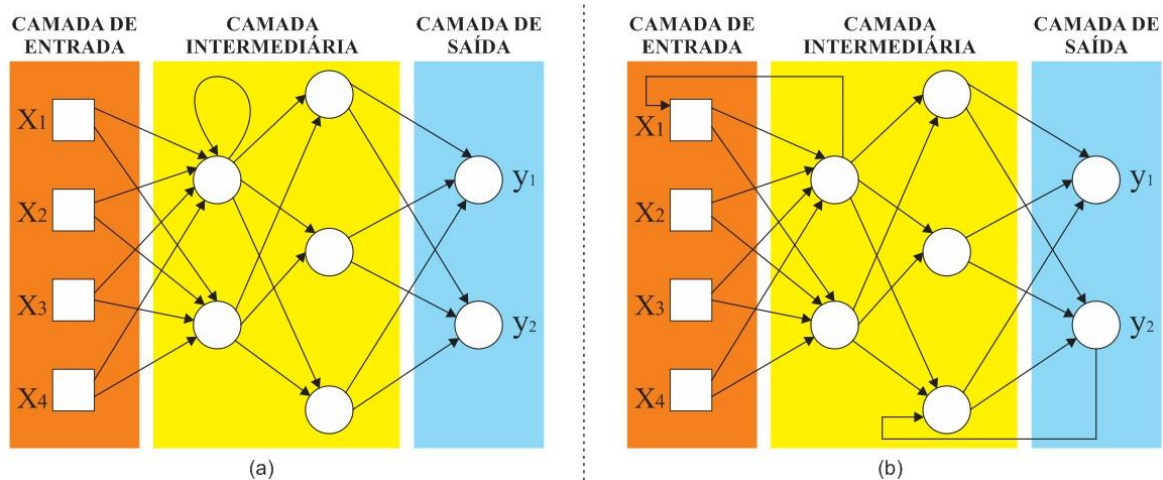


**Figura 15 - Diagrama de uma rede neural artificial multicamada, com duas camadas intermediárias**  
Fonte: Adaptado de COSTA, 2015.

Os tópicos anteriores citados são os principais tipos de conexões dos neurônios *sem-recorrência*. Nas redes *sem-recorrência* (também conhecidas como *feedforward*, diretas ou acíclicas), o sinal percorre a rede em uma única direção, da camada de entrada para a de saída, sendo que os neurônios da mesma camada não são conectados entre si. Em outras palavras, neurônios da  $i$ -ésima camada não podem emitir sinal para neurônios em camadas com índice menores ou iguais a  $i$ . Exemplo de rede *sem-recorrência* é vista na Figura 15.

**Redes recorrentes:** as redes recorrentes (*feedback* ou cíclicas) são aquelas cujo grafo de conectividade possui ao menos um ciclo, ou seja, a saída de pelo menos um neurônio na  $i$ -ésima camada se liga ao menos a um neurônio de uma camada de índice menor ou igual a  $i$  (HASSAN *et al.*, 2016), (COSTA, 2015) e (HAYKIN, 2001), como mostra a Figura 16 (a) e (b). A característica da realimentação qualifica estas redes para processamento dinâmico, isto é, podem ser usadas em sistemas que variam no tempo e necessitam de memória (MARTINIANO *et al.*, 2016).





**Figura 16 - Diagrama de duas redes recorrentes diferentes**  
**Fonte: Adaptado de COSTA, 2015.**

### 6.3 APRENDIZADO

Uma propriedade poderosa de uma RNA está na habilidade de aprender por meio de exemplos, realizar interpolações e extrapolações do aprendizado adquirido e, especialmente, melhorar o seu desempenho pela aprendizagem. A capacidade de adquirir conhecimento é uma das mais importantes qualidades dessas estruturas. Esse é um processo iterativo de ajustes aplicados a seus pesos sinápticos (ANOCHI, 2015).

De acordo com (LARRAIN *et al.*, 2014) no contexto das RNAs, o processo de aprendizagem pode ser definido como um conjunto de regras bem definidas para determinar a intensidade das conexões entre os neurônios, produzindo a solução de um determinado problema de aprendizado. Em geral, uma RNA passa por um período de aprendizado (treinamento), no qual os pesos da rede são ajustados de forma que ela se adapte aos diferentes estímulos que recebe. Para isso, precisa extrair informações e padrões dos estímulos fornecidos a ela (HASSAN *et al.*, 2016). Um algoritmo que descreve uma sequência definida de passos por meio do qual a RNA aprende alguma tarefa é chamado de algoritmo de aprendizado. O tipo de aprendizado depende da maneira com a qual são realizados os ajustes dos parâmetros da rede (MORAIS, 2010).

Diversos métodos para treinamento de redes neurais foram desenvolvidos, podendo ser agrupados em dois principais paradigmas de aprendizagem: aprendizado supervisionado e aprendizado não supervisionado (HAYKIN, 2001; ITURRUGA *et al.*, 2015).

- **Aprendizado supervisionado:** este é o método de aprendizado mais comum no treinamento das redes neurais (ANOCHI, 2015). A resposta desejada é fornecida por um supervisor externo. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos. A RNA tem sua saída calculada comparada com a resposta desejada, recebendo informações do supervisor sobre o erro da resposta atual. Os pesos das conexões são ajustados para minimizar o erro. O exemplo mais popular de algoritmos de aprendizado supervisionado é a regra delta ou de correção de erros. Para a rede de múltiplas camadas, o algoritmo de aprendizagem utilizado é o retropropagação de erros ou *backpropagation* (RUSSEL & NORVIG, 2013; HAYKIN, 2001; RATH *et al.*, 2015; SOARES, 2014).
- **Aprendizado não supervisionado:** para este modelo de aprendizado não há um supervisor (resposta desejada) para indicar a saída desejada aos padrões de entrada. É realizado apenas com os padrões de entrada disponíveis para rede, ao contrário do aprendizado supervisionado. Algoritmos de aprendizagem como o *Hebbiano* e *Competitivo* são exemplos de algoritmos não supervisionados e são especificados em Grossberg (1998 *apud* ANOCHI, 2015).

Neste trabalho, foi utilizado o aprendizado supervisionado, apresentando o algoritmo de treinamento retropropagação do erro (*backpropagation*), o qual é a alternativa mais utilizada para treinamento de redes neurais do tipo *perceptron* de múltiplas camadas (ANOCHI, 2015).

## 6.4 TREINAMENTO

Para que os métodos e tarefas de aprendizagem abordados anteriormente funcionem adequadamente, é necessário que a rede seja submetida a um treinamento. O treinamento é fundamental para qualquer modelo de RNA (HAYKIN, 2001; MARTINIANO *et al.*, 2016; HASSAN *et al.*, 2016). Os principais conceitos relacionados ao treinamento dizem respeito à natureza e à quantidade de dados (MENEZES *et al.*, 2015). Ambos podem prejudicar a generalização e aplicação do modelo. Assim como ocorre na inferência estatística, o

tamanho e a qualidade de uma amostra selecionada têm de estar de acordo com a necessidade do pesquisador. Caso contrário, o processo inferencial pode perder seu poder de generalização ou até ser invalidado (ARRUDA *et al.*, 2013).

Usualmente, o conjunto de dados é dividido em três subconjuntos: o primeiro, nomeado de treinamento, é composto por exemplos que serão utilizados na fase de aprendizagem (treinamento) da rede neural; o segundo, subconjunto chamado de validação, é composto por exemplos que serão utilizados na tarefa de verificar a capacidade de generalização da rede; e, finalmente, o conjunto de teste utilizado para testar o comportamento da rede neural com dados desconhecidos, ou seja, dados que nunca foram vistos antes pela rede neural (ANOCHI, 2015).

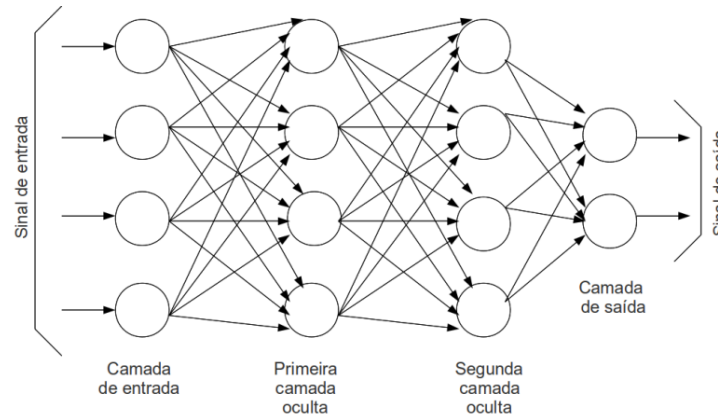
A determinação dos pesos de conexão está ligada à identificação da topologia da rede neural. O algoritmo usado para ajustar os pesos de conexões em uma rede surgiu pela primeira vez em um processo de aprendizagem desenvolvido por Rosenblatt em 1958 para o modelo *perceptron* (SOARES *et al.*, 2014; HAYKIN, 2001).

O treinamento consiste na apresentação dos exemplos do conjunto de treinamento ao sistema. A quantidade de neurônios na camada de entrada é concedida pelo problema a ser tratado. No entanto, a quantidade de neurônios na camada intermediária é característica do projeto. Quando a rede neural possui um alto número de neurônios na camada intermediária, aumenta-se a capacidade de mapeamento não linear da rede. Entretanto, quanto maior esse número, o modelo pode sobre-ajustar aos dados, ocasionando um *overfitting*, dificultando sua capacidade de generalização. Por outro lado, uma rede com poucos neurônios na camada intermediária pode não ser capaz de realizar o mapeamento desejado, o que é denominado de *underfitting* (HAYKIN, 2001; ITURRUGA & SANZ, 2015; KRISTJANPOLLER & MINUTOLO, 2015; LABOISSIERE *et al.*, 2015).

## **6.5 PERCEPTRON DE MÚLTIPLAS CAMADAS - MLP**

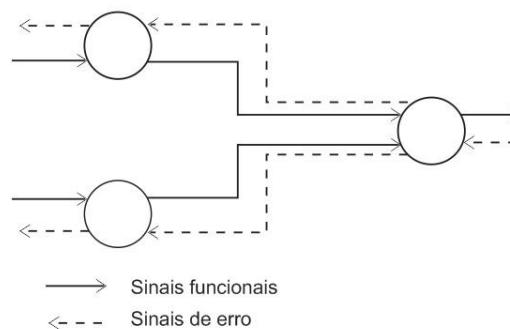
Com a definição do método de treinamento, é necessário optar por algum modelo de RNA. O modelo de RNA utilizado neste trabalho é o modelo *perceptrons* de múltiplas camadas. O modelo *perceptrons* de múltiplas camadas (MLP, do original em inglês *Multilayer Perceptron*) é uma evolução do *perceptron* desenvolvido por Rosenblatt. Redes MLP apresentam um poder computacional muito maior do que aquele apresentado pelas

redes sem camadas intermediárias. Ao contrário destas redes, MLPs podem tratar com dados que não são linearmente separáveis. Teoricamente, redes com duas camadas intermediárias podem implementar qualquer função, seja ela linearmente separável ou não (RATH *et al.*, 2015; RUSSEL & NORVIG, 2013; HAYKIN, 2001). A Figura 17 apresenta a arquitetura básica de um MLP: um modelo com quatro neurônios na camada de entrada, duas camadas ocultas, sendo as duas com quatro neurônios e tendo dois neurônios na camada de saída.



**Figura 17 - Arquitetura de um modelo MLP 4-4-4-2**  
**Fonte: Adaptado de HAYKIN, (2001).**

A rede apresentada na Figura 17 trata de uma rede totalmente conectada, conforme visto anteriormente. As setas representam apenas os sinais de entrada, ou sinais funcionais da rede. Entretanto, existem ainda os fluxos de sinais de erro, que transmitem o valor do erro encontrado, ajustando os pesos sinápticos. A Figura 18 apresenta ambos os fluxos em uma rede.



**Figura 18 - Fluxo de sinais nos neurônios de um modelo MLP**  
**Fonte: Adaptado de Flores, (2009).**

O sinal de entrada parte da camada de entrada, ou camada externa, e vai de camada em camada até, finalmente, a camada de saída, onde o sinal se transforma no resultado da

rede, também chamado de sinal de saída. Nesse momento, a resposta da rede é comparada a uma resposta conhecida, proveniente dos dados de treinamento. Essa comparação gera um erro, chamado aqui de sinal de erro. Este sinal então segue o caminho inverso do sinal de entrada. Durante este percurso é que os pesos sinápticos são ajustados a fim de minimizar esse sinal de erro. Esse algoritmo descrito é o algoritmo de *retropropagação do erro* (do original em inglês, *error back-propagation*) (MARTINIANO *et al.*, 2016; FLORES, 2009).

De acordo com Hassan *et al.* (2016), os modelos de MLP costumam ser confundidos como os únicos modelos de RNA, dadas as suas características e propriedades. Quanto às suas limitações, estas são caracterizadas, principalmente, quanto à não interpretação dos modelos (efeito caixa-preta) e também quanto ao seu uso em amostras pequenas. O nome do efeito caixa-preta se deve à relação entre as caixas-pretas usadas na aviação e o suposto desconhecimento dos seus componentes. Assim como as caixas-pretas, as camadas ocultas de um modelo MLP são conhecidas, mas o seu exato funcionamento, supostamente, não. De tal maneira que os modelos não são interpretáveis quanto aos seus parâmetros, os pesos sinápticos. Contudo, o problema da caixa-preta permanece. O foco principal em uma modelagem por MLP é obter previsões e não o conhecimento da função geradora ou de seus parâmetros (COSTA, 2015).

## 7 – METODOLOGIA

---

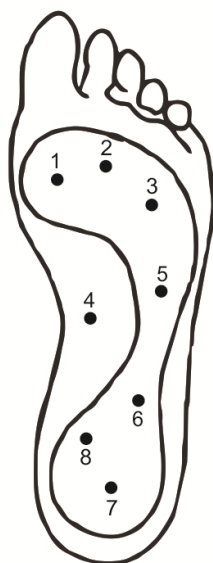
Neste capítulo, os métodos implementados na pesquisa são apresentados. Foram utilizados os dados experimentais de Ximenes & Peron (2013) e aplicado o método da Dinâmica Simbólica otimizada por Algoritmo Genético para análise e comparação dos dados.

### 7.1 DADOS DA PRESSÃO PLANTAR DA MARCHA

Os problemas do sistema motor e a dificuldade de comunicação dos indivíduos autistas inspiraram a realização desse método que se iniciou com a utilização dos dados de Pressão Plantar da Marcha (PPM), oriundos do trabalho experimental de Ximenes & Peron (2013), onde apresentaram e desenvolveram uma ferramenta que se propôs a auxiliar a identificação de alterações na pisada de crianças autistas por meio da coleta e avaliação da distribuição da Pressão Plantar em comparação com o Grupo de Controle (GCO). Trata-se de uma palmilha individualizada e personalizada de látex, com sensores estrategicamente posicionados, de forma a coletar os sinais de interesse.

Existem algumas técnicas para captura dos sinais obtidos na palmilha; a escolhida, segundo Santarmou (2006 *apud* XIMENES & PERON, 2013), tem por base a utilização de sensores de forma isolada, permitindo uma medição de pressão local, ou em grupo sob a forma de matriz. Essa medição tem como vantagens ser de fácil uso e possuir um custo de aquisição reduzido, mesmo para uma frequência de amostragem elevada, devido ao número de sensores utilizados ser pequeno SOUSA *et al.* (2006 *apud* XIMENES & PERON, 2013).

Na palmilha, os sensores estão fixados com fita adesiva para garantir que não ocorra deslocamento deles durante os testes. Foram distribuídos 8 sensores ao longo da palmilha conforme apresenta a Figura 19, sendo os pontos de 1 a 3 pertencentes ao salto do pé, 4 e 5 ao médio pé e de 6 a 8 ao metatarso. Tais regiões, com exceção do ponto 4, foram escolhidas por suportarem a maior parte do peso do corpo; portanto, a distribuição dos sensores segundo esse layout permite a identificação da pressão plantar nos pontos de maior aplicação de força. O ponto 4 permite identificar se o pé estudado apresenta caracterização de um pé plano, onde toda a sola do pé entra em contato com o chão.



**Figura 19 - Distribuição dos sensores na palmilha**  
**Fonte: Ximenes & Peron, 2013.**

Os resultados do modelo de captação de dados de PPM de Ximenes & Peron (2013) foram bastante importantes, pois analisaram pontos estratégicos da região plantar do pé, tendo sido captadas duas amostras de PPM do GAU e três amostras de PPM para o GCO.

## **7.2 FLUXOGRAMAS DAS ETAPAS DA METODOLOGIA**

O fluxograma das etapas da metodologia proposta está dividido em três partes compostas por: Otimização da dinâmica simbólica por algoritmo genético; Avaliação do desempenho da metodologia desenvolvida; Identificação automática da presença de autismo pelas séries de pressão plantar da marcha, através de redes neurais artificiais

Cada uma delas descrita como uma sequência de fases, conforme apresentado nos tópicos seguintes.

### 7.2.1 OTIMIZAÇÃO DA DINÂMICA SIMBÓLICA POR ALGORITMO GENÉTICO

A primeira etapa conforme Figura 20, inicia-se com aquisição dos dados de PPM (Fase 1), posteriormente estes dados são estruturados e cadastrados na plataforma de Banco de Dados (BD) desenvolvida para armazenar as informações de entrada/processamento/saída dos dados da aplicação desta metodologia (Fase 2). O banco de dados foi desenvolvido utilizando-se o Sistema Gerenciador de Banco de Dados (SGBD).

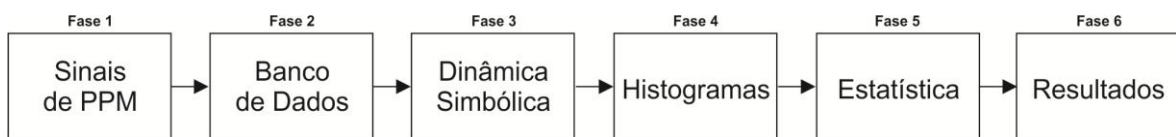


**Figura 20 - Fluxograma da Otimização dos parâmetros da Dinâmica Simbólica por Algoritmo Genético.**

Em seguida, na (Fase 3), executa-se o AG concomitantemente com a DS, com o objetivo de se otimizar os parâmetros da DS (*tamanho do alfabeto e tamanho da palavra*). O fluxograma é finalizado na Fase 4 com os resultados obtidos pela otimização. Esses resultados serão utilizados nas etapas de Avaliação do Desempenho da Metodologia (item 7.2.2) e na Identificação Automática da Presença do Autismo (item 7.2.3).

### 7.2.2 AVALIAÇÃO DO DESEMPENHO DA METODOLOGIA DESENVOLVIDA

Conforme a Figura 21, segue-se com as mesmas aplicações descritas no item 7.2.1 para as Fases 1 e 2. Na Fase 3 é aplicada a DS para simbolização das séries temporais utilizando os valores dos parâmetros otimizados na primeira etapa descrita no item 7.2.1.



**Figura 21 - Fluxograma para avaliação do desempenho da metodologia.**

A distribuição das frequências foi verificada por meio de histogramas, resultantes da DS (Fase 4). Prosseguindo para a Fase 5, podemos através de avaliações estatísticas, testar as diferenças entre as séries temporais da PPMs do GAU e do GCO. Finalizando com a Fase 6, são mostrados os resultados que possam ser analisados afim de realizar comparações da metodologia.



### 7.2.3 IDENTIFICAÇÃO AUTOMÁTICA DA PRESENÇA DE AUTISMO PELAS SÉRIES DE PRESSÃO PLANTAR DA MARCHA POR MEIO DE REDES NEURAIAS ARTIFICIAIS

As Fases 1 a 4 da Figura 22 seguem a mesma metodologia descrita no item 7.2.2. Em seguida na Fase 5 são calculados os valores das entropias de cada histograma (Fase 4) do sinal de cada eletrodo. Estes valores serão, posteriormente, fornecidos para a entrada da RNA.



Figura 22 - Fluxograma das fases para classificação automática dos sinais de PPM.

Na Fase 6 e 7, com o treinamento prévio da RNA, são executados os testes, a fim de demonstrar a capacidade do método em gerar resultados de forma automática e verificar a sua eficiência em diferenciar as estruturas simbólicas dos dados de PPM do GAU e PPM do GCO encontrados após a simbolização das séries temporais, onde também é analisado estatisticamente seu desempenho.

## 7.3 ESTRUTURAÇÃO DOS DADOS

Como visto na Figura 20 - Fase 2, executa-se a estruturação e planejamento das tabelas do banco de dados, proporcionando um ganho no tempo de processamento, realização de buscas, seleção de informações, ajustes de experimentos, segurança e integridade dos dados, pois o grande fluxo de processamento executado pelo AG e a DS poderá demandar tempo, permitir a perda ou o dano nos dados. A Figura 23 detalha o modelo de relacionamento das tabelas do Banco de Dados que foram projetadas para armazenar e processar todas as informações necessárias para o método.

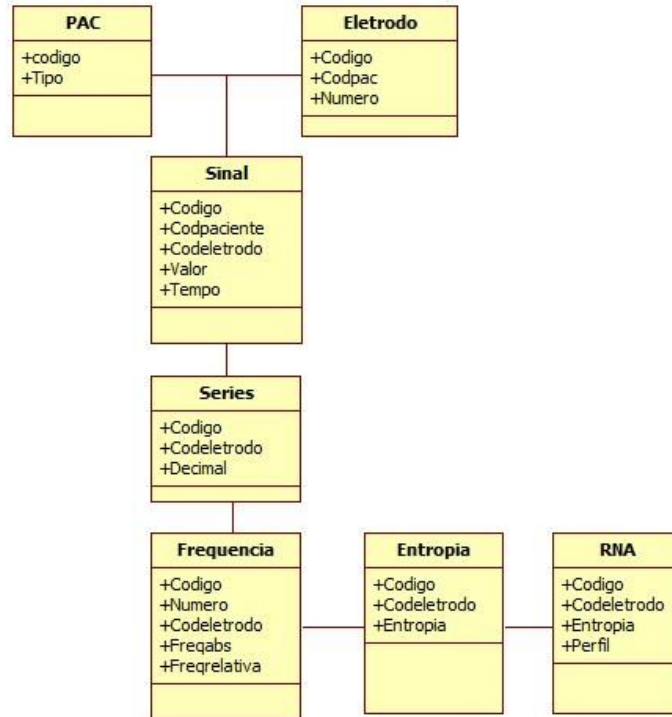


Figura 23 - Relacionamento das Tabelas do Banco de Dados

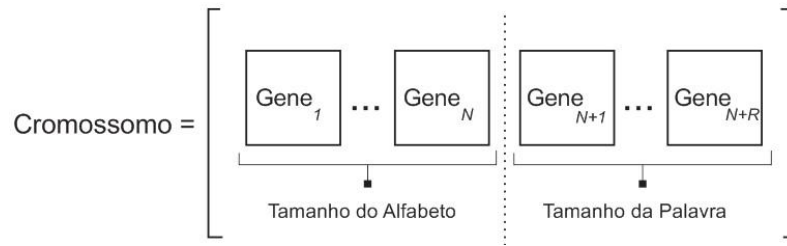
## 7.4 OTIMIZAÇÃO DOS PARÂMETROS DA DINÂMICA SIMBÓLICA POR ALGORITMO GENÉTICO

Conforme Figura 20 – Fase 3, o AG foi implementado dentro da rotina computacional e desenvolvido com o uso do ambiente de programação Delphi 7. As etapas para o desenvolvimento da metodologia empregada com a utilização dos AGs estão descritas nos itens subsequentes.

### 7.4.1 CODIFICAÇÃO DOS INDIVÍDUOS E POPULAÇÃO INICIAL

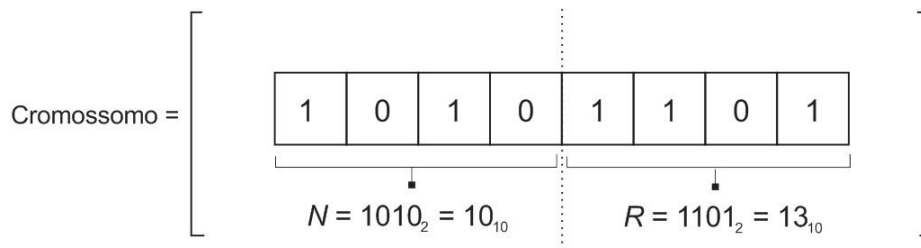
Inicia-se o processo de otimização com a definição da estrutura dos indivíduos que irão compor as populações que serão geradas. Essa estrutura é denominada *Representação Cromossômica*, neste trabalho, cada cromossomo (solução) é composto pela justaposição de dois parâmetros inteiros representados por valores binários ou por  $(N + R)$  genes representados por valores binários, onde  $N$  representa o tamanho do *Alfabeto* e  $R$ , o tamanho da *Palavra*, ambos representando os parâmetros da DS. Os primeiros  $N$  genes serão usados para serem codificados definindo a quantidade de partições (Tamanho do Alfabeto) da DS e os últimos genes representarão o número de símbolos agrupados (Tamanho da Palavra), ambos serão posteriormente codificados em números decimais. A

Figura 24 ilustra a estrutura do cromossomo e o exemplo de codificação, pode ser visto na Figura 25.



**Figura 24 - Estrutura do Cromossomo codificado para o método**

Na terminologia dos AGs, cada indivíduo ou cromossomo representa uma solução. Para isso cada grupo de gene é convertido de seu valor binário para decimal e esse valor equivale a uma alternativa para os parâmetros da simbolização, conforme ilustrado na Figura 25:



**Figura 25 - Cromossomo codificado em duas sessões com sua representação decimal**

A população inicial (geração um) do AG é composta de 100 indivíduos escolhidos aleatoriamente.

#### 7.4.2 FUNÇÃO DE APTIDÃO (AVALIAÇÃO)

O propósito da otimização realizada pelo AG é maximizar ou minimizar uma função de aptidão (também chamada de adequação, custo, objetivo ou *fitness*). Sendo assim, um dos aspectos mais importantes de um AG é a escolha de uma função de aptidão apropriada, pois será utilizada para medir o desempenho dos indivíduos. Essa função avalia cada indivíduo de cada geração, atribuindo-lhes uma nota, que reflete sua aptidão em resolver o problema estudado. Então, nos modelos utilizados para representar os problemas de simbolização das séries temporais, a aptidão de cada indivíduo é obtida por meio da Equação 1:

$$f(x) = \left| \sum_k \sigma_k - \sum_j \sigma_j \right|, \quad \text{Equação 1}$$

em que  $f(x)$  é a função de aptidão utilizada para representar a aptidão dos indivíduos da população;  $\sigma_k$  = entropia dos indivíduos do GCO;  $\sigma_j$  = entropia dos indivíduos do GAU. A entropia é definida por meio da Equação 2:

$$\sigma = -\frac{1}{\log N_{obs}} \sum_i p_{i,L} \log p_{i,L}, \quad \text{Equação 2}$$

onde  $\sigma$  = entropia de Shannon;  $N_{obs}$  = número de colunas do histograma com valor diferente de zero;  $i$  = índice da sequência de símbolos do vetor  $L$ ; e  $p_{i,L}$  = frequência relativa dos valores dos índices  $i$  de cada coluna  $L$  do histograma.

A função de aptidão é aplicada a todos os indivíduos dos dois grupos, utilizando o algoritmo da Dinâmica Simbólica e quantificando os resultados por meio das Equação 1 e Equação 2.

A *Entropia de Shannon* é a medida de incerteza de uma variável aleatória. Na Teoria da Informação a entropia é aplicada em pesquisas no âmbito de técnicas de reconhecimento de padrões. Mais especificamente, refere-se ao uso da entropia como medida capaz de fornecer informação sobre variáveis que melhor discriminam os dados, avaliando a relativa complexidade das distribuições das sequências de símbolos. (SANTOS, 2013), (PINEDA, 2010) e (DAW *et al.*, 2003); portanto, comparando os sinais de PPM, quanto maior a diferença do grau de incerteza entre os dados PPM do GAU e GCO maiores serão as possibilidades de existência de estrutura simbólicas que diferenciem esses dados, conseqüentemente, o indivíduo que obtiver o maior valor para a função objetivo, que representa diferença em módulo das entropias do GAU e do GCO, representa o conjunto de parâmetros da simbolização que melhor distingue os grupos analisados.

A diferenciação dos graus de incerteza encontrados por meio do resultado da Equação 1 será o valor da Aptidão do indivíduo em sua avaliação.

### 7.4.3 SELEÇÃO DOS INDIVÍDUOS

A Seleção consiste em escolher os indivíduos da população que irão se reproduzir (cromossomos-pais). Com base em sua aptidão, os indivíduos são selecionados por elitismo. Por elitismo, cinco indivíduos mais aptos (TOPO) são selecionados para passarem para a geração seguinte sem alterações. A vantagem dessa estratégia é garantir que as melhores soluções sejam repassadas para as próximas gerações.

Para garantir a evolução dos indivíduos, aplicam-se os Operadores Genéticos, criando, a partir de uma população, novas populações ou gerações de indivíduos melhorados. A cada nova população ou geração espera-se melhoria no desempenho dos indivíduos em relação ao meio ambiente, o que traduzido para o processo de busca de novas soluções pelo AG, neste trabalho, significa obter maiores valores para a função objetivo.

### 7.4.4 OPERADORES GENÉTICOS

Utilizam-se como Operadores Genéticos o *Cruzamento* e a *Mutação*. O operador de Cruzamento permite a obtenção de novos indivíduos (filhos), a partir da combinação dos cromossomos dos pais. O tipo de cruzamento escolhido foi de Um-Ponto. No cruzamento de um único ponto, os dois indivíduos pais selecionados para reproduzir têm seus materiais genéticos trocados a partir de um ponto de ruptura, escolhido aleatoriamente. A partir desse cruzamento são gerados dois indivíduos com a combinação do código de cada um dos pais (Figura 26). O cruzamento é aplicado entre um indivíduo que está classificado entre os mais bem avaliados da população, e outro aleatoriamente selecionado. O resultado desse cruzamento será a geração de novos indivíduos (filhos) que farão parte da nova população.

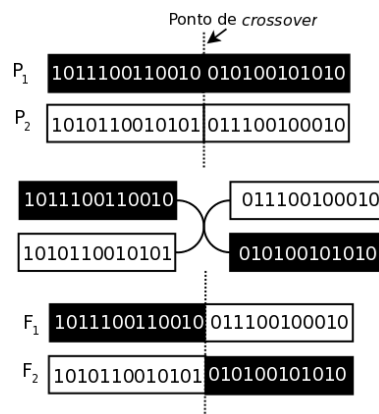
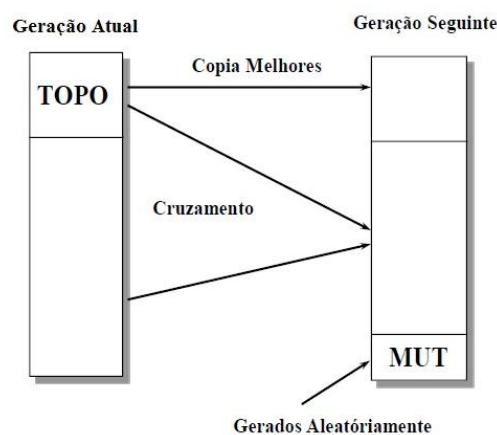


Figura 26 - Crossover de Um-Ponto gerando novos indivíduos

Após a ocorrência do cruzamento, alguns indivíduos da população serão sorteados para aplicação da operação de Mutação. O operador de Mutação altera os valores de alguns bits do cromossomo, conferindo diversidade à população. Porém, ao invés de se realizar a mutação gene a gene, opta-se em adicionar a cada nova geração novos membros a partir da mesma distribuição utilizada para a geração da população inicial (MUT). A rigor, este procedimento pode ser considerado uma mutação radical que não guarda resquício da população anterior. Esse processo visa evitar a convergência prematura de uma população. O valor da taxa de probabilidade de mutação nesse modelo será de 25%. A Figura 27 ilustra o processo de transição de uma geração para a seguinte:



**Figura 27 - Transição de geração**

Esses parâmetros são selecionados em uma fase de testes para determinação de melhores parâmetros do AG, no qual várias combinações foram testadas. Em cada simulação computacional realizada, esse algoritmo é executado 10 vezes, cada vez com um valor aleatório diferente. Esse procedimento é importante porque a população inicial e os operadores de seleção e reprodução são estocásticos.

#### **7.4.5 CRITÉRIO DE PARADA**

Os procedimentos iterativos requerem a definição de um critério de parada para interromper a corrida do algoritmo. Segundo Holland (1975), um dos problemas enfrentados na confecção de um AG é determinar o critério de parada. Essa dificuldade se deve à inabilidade ou dificuldade desses algoritmos em avaliar a qualidade da solução em um dado instante da busca, pois, em geral, não há informação de quanto à solução corrente se aproxima da solução ótima.

Há, contudo, alguns mecanismos para definição do critério de parada. O número máximo de gerações, o tempo máximo de processamento e a estabilização da função objetivo são alguns exemplos. A escolha desse critério pode também afetar a qualidade da solução encontrada. Por exemplo, ao utilizar como critério de parada a primeira solução viável encontrada, uma solução de baixa qualidade pode ser escolhida, uma vez que para um dado problema pode existir uma infinidade de soluções viáveis.

O AG proposto utiliza como critério de parada o *Número Máximo de Gerações*. Nesse modelo, são definidas 1000 gerações com 100 indivíduos, e verificamos as aptidões dos melhores indivíduos da última geração, caso o valor satisfaça nosso objetivo, prosseguiremos para a próxima etapa; caso contrário, repetiremos todo o processo indicando para a população inicial os valores maximizados da última geração do experimento anterior. É importante observar se ocorre convergência no processamento do AG, que implica na não ocorrência de melhoras significativas na solução durante um dado número de gerações (COSTA, 2015). Testes preliminares foram necessários para se determinar um número adequado de gerações para interromper o processo de otimização.

Ao final do processamento do AG, a última população terá o indivíduo com maior aptidão, e seus valores como parâmetros serão os que se aproximam de uma melhor solução para o problema.

## **7.5 SIMBOLIZAÇÃO DAS SÉRIES TEMPORAIS**

De acordo com a Figura 21 – Fase 3, aplica-se a Dinâmica Simbólica. O conceito da dinâmica simbólica (DS) é baseado na seqüência de visitas do sinal a partições do espaço de fases. Dessa forma, cada série de tempo é transformada em uma seqüência de símbolos a partir de um alfabeto. São definidos parâmetros como o *tamanho do alfabeto* ( $n$ ), que é utilizado na transformação das séries e o *tamanho da palavra* ( $m$ ) que é o número de símbolos agrupados para análise das séries. Neste trabalho, esses parâmetros são definidos pelo resultado da otimização por Algoritmo Genético. A otimização destes parâmetros é aplicada para se obter medidas capazes de diferenciar os dois grupos (GAU e GCO).

Para caracterização e análise dos dados de PPM, é utilizado o método de simbolização de séries temporais experimentais descrito por Souza (2012), Baumert *et al.*

(2015), Valencia *et al.* (2015), Kabir *et al.* (2011), Daw *et al.* (2003) e Daw *et al.* (2002). A DS é uma abordagem não linear com base nas transformações de séries em sequências simbólicas com símbolos de um determinado alfabeto, para investigar os aspectos dinâmicos de uma série temporal. Ela fornece informações simplificadas de sistemas complexos, por meio de um número limitado de símbolos (VALENCIA *et al.*, 2015). Algumas informações detalhadas são perdidas nesse processo, mas a sequência simbólica obtida ainda descreve de forma adequada o comportamento dinâmico da série original.

As transformações simbólicas dos dados aplicadas neste estudo são definidas entre Estática e Dinâmica. Na transformação simbólica estática são atribuídos para uma determinada medida  $n$  valores simbólicos binários (por exemplo, 0 ou 1 para  $n=2$ ), em que o limiar de particionamento  $\ell$  é definido por:

$$\ell = \frac{1}{N} \sum_{i=1}^N x_i, \quad \text{Equação 3}$$

onde  $\ell$  é média aritmética simples com  $N$  representando o número total de dados de um sinal e  $x_i$  o valor do sinal em  $i$ . Nessa abordagem tendo em conta a série temporal  $x=\{x(i), i = 1, \dots, N\}$  a série completa  $x(i)$  está dividida em  $N$  partes quantificadas pelo limiar  $\ell$ , a fim de obter uma série simbólica  $S_i$ , tal como é definido pela seguinte Equação 4, que será transformada em símbolos  $A = \{1;0\}$ .

$$S_i = \begin{cases} 1: x_i \geq \ell \\ 0: x_i < \ell \end{cases} \quad \text{Equação 4}$$

Na abordagem da transformação simbólica dinâmica, a série  $x(i), i = 1, \dots, N$  os valores obtidos para a sequência de símbolo  $S_i$  são determinados pela posição de  $N$  conforme Equação 7. Os valores de  $n$  são representados de acordo com um valor gradual (por exemplo, 0, 1, 2 para  $n = 3$ ), nesse caso o limiar de particionamento  $\ell$  será um vetor  $\ell=\{\ell(i), i = 1, \dots, n-1\}$  definido por:

$$\ell_{(i)} = \min(x(i)) + \sum_{i=1}^{n-1} \mu, \quad \text{Equação 5}$$



onde  $\min(x(i))$  é o menor valor da série  $x(i)$ ; o vetor da função  $\ell(i)$  é o valor do limiar de particionamento em  $i$ . O valor de  $\mu$  é determinado por:

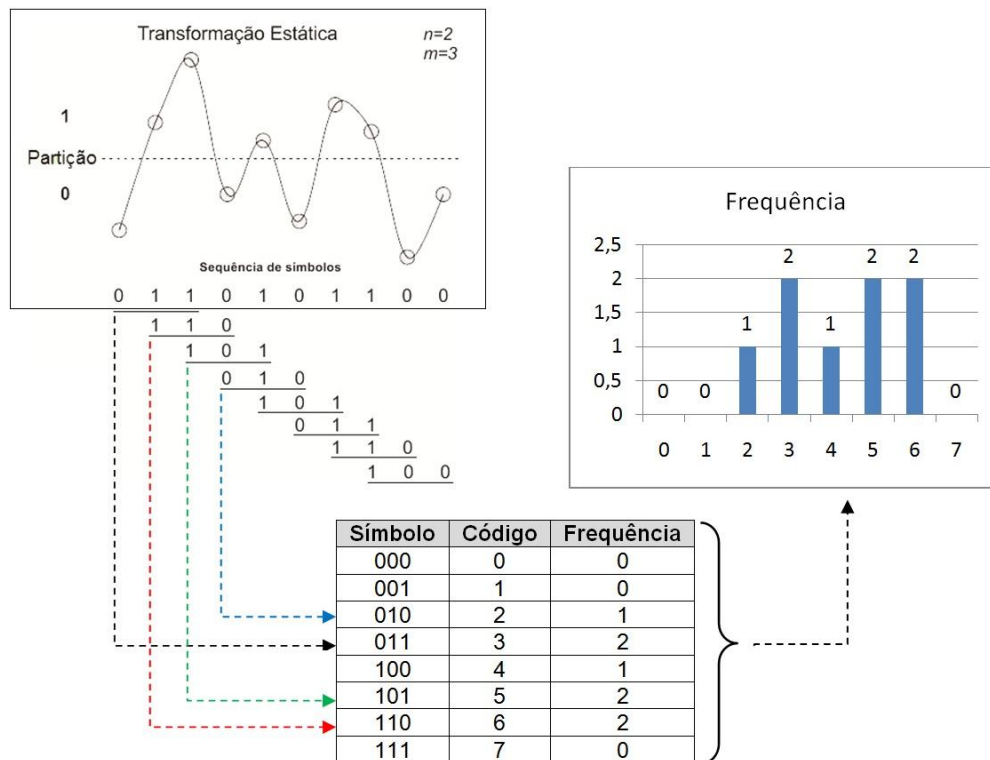
$$\mu = \frac{\max(x(i)) + |\min(x(i))|}{n - 1}, \quad \text{Equação 6}$$

onde  $\max(x(i))$  é o maior valor da série  $x(i)$  e  $n$  é o valor da quantidade de símbolos em que a série  $x(i)$  será transformada para formar a série simbólica  $S_i$  em símbolos  $A=\{0, 1, \dots, n\}$ , como definido pela Equação 7:

$$S_i = \left[ \begin{array}{l} N: \quad x_i \geq \ell_n \\ \vdots \\ 2: \quad \ell_2 \leq x_i < \ell_3 \\ 1: \quad \ell_1 \leq x_i < \ell_2 \\ 0: \quad x_i < \ell_1 \end{array} \right], \quad \text{Equação 7}$$

Uma vez que uma série temporal é simbolizada, serão avaliadas as frequências relativas de todas as possíveis sequências de símbolos definidas por um vetor de comprimento  $m$ , que são formados utilizando uma abordagem de janela deslizante, onde  $A=[a_n, a_{n+1}, \dots, a_{N-(m-1)}]$ . Por exemplo, se  $m=3$ , a frequência relativa de ocorrência para cada possível sequência de combinação de 3 símbolos (“palavras de 3 letras”) é computada.

Uma maneira conveniente de analisar as frequências das séries simbólicas é atribuir um número único para cada sequência possível atribuindo seu valor decimal equivalente, sendo definido um número decimal como código de sequência. Por exemplo, uma sequência  $101$  ocorrendo com  $n=2$  teria um código de sequência de valor 5. Considerando o esquema de simbolização binária, palavras de comprimento de  $m=3$  terão 8 tipos diferentes de palavras ( $2^3 = 8$ ). Após esta codificação, para cada sequência de símbolos, as frequências relativas são representadas por meio de histogramas de frequência de ocorrência. Esses histogramas são utilizados para análise estatística. A Figura 28, exemplificando o modelo de aplicação utilizado, demonstra sinteticamente o método da DS até a obtenção das frequências absolutas, que posteriormente serão normalizadas para se obter as frequências relativas.



**Figura 28 – Ilustração sintética do método da Dinâmica Simbólica com Transformação Simbólica Estática**

A partir da discussão acima, observa-se que as sequências de símbolos serão utilizadas para representar quaisquer possíveis variações ao longo do tempo. Essa é uma propriedade poderosa, por não fazer qualquer suposição sobre a natureza dos padrões funcionando igualmente para fenômenos lineares, não lineares e aleatórios (VALENCIA *et al.*, 2015).

## 7.6 IDENTIFICAÇÃO AUTOMÁTICA DA PRESENÇA DE AUTISMO

Conforme a Figura 22 – Fase 6, aplica-se a RNA para classificação automática das séries de PPM. Foram utilizadas Redes Neurais Artificiais (RNA) do tipo Multicamadas (MLP) para classificar os sinais de PPM pela Dinâmica Simbólica. Entre vários motivos pela escolha do modelo MLP, citam-se as características deste modelo:

- Não linear capaz de modelar funções complexas;
- Robusto, sofrendo pouca influência de ruídos;
- Adaptável aos pesos e topologias com base nas mudanças de ambiente;

- Facilidade de uso, podendo ser usado com pouco conhecimento sobre o relacionamento das funções a serem modeladas (visto como uma “caixa-preta”).

Mesmo com várias características positivas que determinaram a escolha do tipo de RNA, deve-se nos preocupar com algumas falhas a que é suscetível esse modelo. Dentre elas pode-se citar os problemas de generalização do tipo *over-training* (*Sobre-treinamento, quando há um excesso nas iterações causando falhas de treinamento*) e o *overfitting* (*Sobre-ajuste, quando o número de parâmetros é grande e o classificador tende a se adaptar a detalhes específicos da base de treinamento*). Além da possibilidade de ocorrência dessas falhas, pode existir o inconveniente de exigir de um longo tempo de treinamento.

A RNA MLP é constituída pelas camadas de entrada, oculta e de saída. A camada de entrada é definida por oito entradas, que irão receber os valores da *Entropia de Shannon* (ver Equação 2), referente aos valores de cada eletrodo de acordo com suas séries temporais simbolizadas (conforme Figura 22 – Fase 5). A camada oculta irá variar entre 2 a 8 neurônios. Após a realização de alguns testes preliminares, observou-se que o uso de mais de oito neurônios não produzia melhoria significativa nos resultados, apenas aumentava significativamente o tempo de treinamento. Na camada de saída, definimos duas possibilidades para representar o resultado: como pertencendo ao GAU ou ao GCO.

A RNA utilizada nesta pesquisa foi construída, treinada e testada utilizando o componente *TMLP* da Linguagem de Programação *Delphi 7*. Dessa forma, a nomenclatura adotada para as funções de ativação e algoritmos de treinamento é a mesma desenvolvida para o software, o *Backpropagation*.

Na rede MLP existem duas fases distintas denominadas de treinamento/aprendizagem e teste/prova. Para que uma RNA possa fornecer resultados adequados, é necessário que passe por uma fase de treinamento, na qual seus pesos são ajustados de forma que ela se adapte aos diferentes estímulos de entrada. Durante essa fase de treinamento, ocorre o seu aprendizado (BANDEIRA, 2015).

Neste trabalho, é aplicado o *Aprendizado Supervisionado*, em que informamos à RNA uma referência do objetivo a ser alcançado (por exemplo, informando a entrada de

dados com características do GCO indicamos que o resultado correto será o perfil GCO e assim da mesma forma para os dados de perfil GAU). Essas técnicas citadas representam o treinamento da MLP, preparada para que a RNA se torne capaz de se aproximar ao máximo do objetivo descrito, como listado na Tabela 2:

**Tabela 2 – Tabela dos Perfis dos Sinais indicados no treinamento da RNA.**

<i><b>PERFIL</b></i>	<i><b>DESCRIÇÃO</b></i>	<i><b>OBJETIVO</b></i>
<b>GAU</b>	Sinal de PPM do GAU	Informar para camada de saída o Perfil GAU
<b>GCO</b>	Sinal de PPM do GCO	Informar para camada de saída o Perfil GCO

A Figura 29 ilustra sinteticamente o fluxo do processo para classificação automática dos sinais, exemplificando os passos desde a entrada de dados de cada eletrodo no algoritmo da DS até o resultado da classificação pela RNA.

Esta Figura resume esquematicamente os passos para classificação automática por RNA, iniciando com a entrada dos dados de PPM, e, em seguida, é executada a transformação simbólica. Histogramas são gerados com os resultados da simbolização mensurados por meio da Entropia de Shannon (Equação 2). O valor da entropia de cada série simbolizada referente ao eletrodo irá alimentar as entradas da RNA. O resultado da RNA é informado pela sua camada de saída, caracterizando o sinal de PPM como pertencendo a um dos perfis treinados.

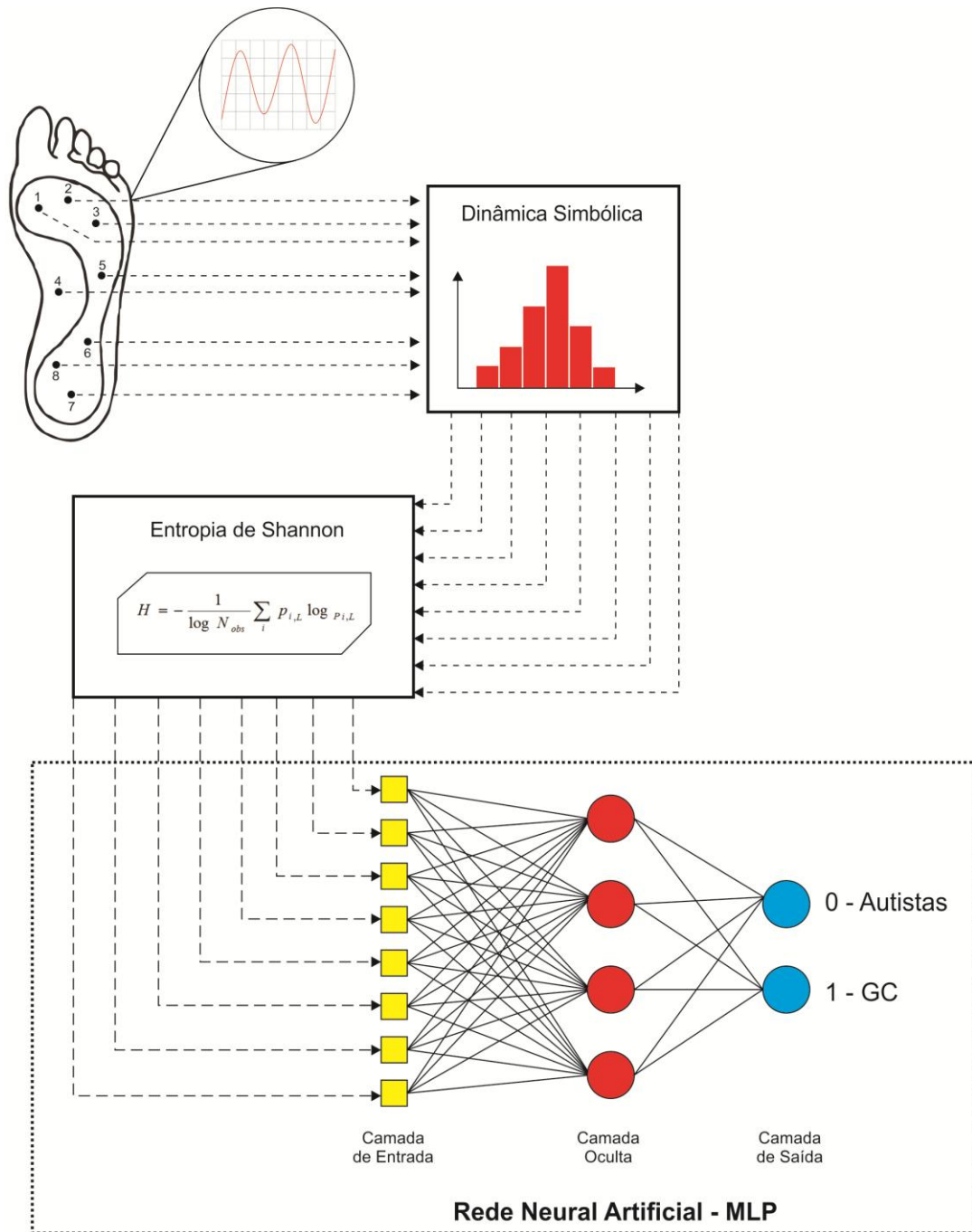


Figura 29 - Metodologia para classificação automática dos sinais

## 7.7 ANÁLISE ESTATÍSTICA DAS SÉRIES SIMBÓLICAS

Conforme Figura 21, na Fase 5, foi utilizado para as análises estatísticas o *software* GraphPad Prism, versão 7 para Windows (*download* em: [www.graphpad.com](http://www.graphpad.com)). Todas as análises consideraram as sequências de símbolos representadas pelos seus respectivos histogramas. Essas representações, além de proporcionar uma visualização dos padrões da dinâmica, fornecem a base para as estatísticas quantitativas. Neste trabalho utilizamos a

*Entropia de Shannon* definida pela Equação 2, a *Norma Euclidiana* e a estatística  $\chi^2$  *modificada*, estas duas últimas úteis para a diferenciação de histogramas (DAW *et al.*, 2002; DAW *et al.*, 1998; DAW *et al.*, 2003). A Norma Euclidiana é definida como:

$$T_{AB} = \sqrt{\sum_i (A_i - B_i)^2}, \quad \text{Equação 8}$$

e a estatística  $\chi^2$  *modificada* é definida por:

$$\chi^2_{AB} = \sum_i \frac{(A_i - B_i)^2}{(A_i + B_i)}, \quad \text{Equação 9}$$

em que  $A_i$  e  $B_i$  são as frequências relativas individuais da sequência  $i$  dos histogramas de  $A$  e  $B$ . Ambas as estatísticas acima são geradas pela diferenciação das frequências das sequências específicas de diferentes conjuntos de dados. Quando as diferenças das frequências são grandes, os indicadores estatísticos resultantes também serão grandes. Da mesma forma, grandes valores para os indicadores estatísticos implicam que os padrões dinâmicos nos conjuntos de dados são diferentes. A Norma Euclidiana baseia-se na idéia de que cada sequência de símbolos do histograma pode ser pensada como um vetor no espaço multidimensional, em que o número de dimensões é o número de possíveis sequências únicas. Assim, os valores das diferenças entre os histogramas devem proporcionar uma comparação entre eles. A estatística  $\chi^2$  *modificada* é derivada a partir da estatística  $\chi^2$  *padrão*, em que as frequências de medições univariadas foram substituídas por frequências das sequências simbólicas (DAW *et al.*, 2002).

As frequências de ocorrência de palavras, observadas por meio dos histogramas foi comparadas também pela análise de medidas não paramétrica de variância (medidas descritivas, teste *t Student* e teste de Qui-quadrado). Valores com  $p \leq 0,05$  foram considerados estatisticamente significativos. Os dados foram expressos como média e desvio padrão.

## 7.8 ANÁLISE ESTATÍSTICA DO DESEMPENHO DA RNA

De acordo com Figura 22 – Fase 7, foram utilizadas para análise estatística do desempenho da RNA as avaliações descritas por Subasia *et al.* (2005), representadas de acordo com as medidas estatísticas especificadas abaixo:

- VERDADEIROS POSITIVOS ( $\alpha$ ): eventos que foram identificados pelo sistema como Características do GAU e confirmados pela marcação de um especialista.
- FALSOS NEGATIVOS ( $\beta$ ): eventos que foram identificados pelo sistema como não sendo do GAU, porém, os especialistas marcaram como sendo do GAU.
- VERDADEIROS NEGATIVOS ( $\gamma$ ): eventos que foram identificados pelo sistema como não sendo do GAU e confirmados pelos especialistas.
- FALSOS POSITIVOS ( $\delta$ ): eventos que foram identificados pelo sistema como sendo do GAU, mas não sendo confirmados pelos especialistas.

As grandezas que medem o desempenho da RNA em expressar a relação entre os dados de PPM do GAU e GCO são:

$$\pi = \frac{\alpha}{\alpha + \beta} \times 100[\%], \quad \text{Equação 10}$$

onde  $\pi$  é a Sensibilidade (ou Relação Verdadeira Positiva) que avalia a precisão em detectar corretamente as características de PPM do GAU.

$$\varphi = \frac{\gamma}{\gamma + \delta} \times 100[\%], \quad \text{Equação 11}$$

onde  $\varphi$  é a Especificidade (ou Relação Verdadeira Negativa) que avalia a precisão em detectar características de PPM do GCO.

$$\varepsilon = \frac{\alpha}{\alpha + \delta} \times 100[\%], \quad \text{Equação 12}$$

onde  $\varepsilon$  é a Seletividade Positiva (ou Valor Preditivo Positivo), representa a capacidade em indicar corretamente as PPM do GAU.

$$\tau = \frac{\gamma}{\gamma + \beta} \times 100[\%], \quad \text{Equação 13}$$

onde  $\tau$  é a Seletividade Negativa (ou Valor Preditivo Negativo), corresponde a capacidade em indicar corretamente as PPMs do GCO.

Para estabelecer a capacidade do sistema em reconhecer os dados, foi aplicado o quantificador (Equação 14) para reconhecer o sucesso do sistema, denominado Precisão (SUBASIA, *et al*, 2005).

$$\theta = \frac{\pi + \varphi}{2}, \quad \text{Equação 14}$$

onde  $\theta$  é a função do quantificador de Precisão,  $\pi$  é a sensibilidade definida pela Equação 10 e  $\varphi$  é a especificidade definida pela Equação 11.



## 8 – RESULTADOS E DISCUSSÕES

---

Neste capítulo, são apresentados os resultados obtidos pelos métodos descritos no capítulo anterior, referente à metodologia, com suas respectivas descrições e discussões.

### 8.1 OTIMIZAÇÃO DE PARÂMETROS

Aplicou-se a otimização por Algoritmo Genético aos dados de PPM do GAU e do GCO, com o objetivo de encontrar valores ótimos de parâmetros a serem utilizados na simbolização das séries temporais. Foi necessária a otimização de dois parâmetros (*tamanho da palavra* e *tamanho do alfabeto*) para simbolização dessas séries, a fim de se evitar o arbítrio de parâmetros e excessivos testes experimentais. Inicia-se a otimização, com a aplicação dos valores descritos na Tabela 3 para cada operador genético responsável pela configuração e atuação dos dados dentro do AG.

Tabela 3 – Configuração dos operadores genéticos utilizados na otimização.

<i>OPERADORES GENÉTICOS</i>	<i>VALORES</i>
Tamanho do cromossomo	8
Tamanho da população	100
Taxa de crossover	70%
Taxa de mutação	25%
Pontos de crossover	1
Indivíduos no elitismo	5
Número de gerações	1000
Tipo de seleção	Elitismo
Tipo de otimização	Maximização

Definiu-se o *tamanho do cromossomo* de acordo com o tamanho máximo da *palavra* e do *alfabeto*, orientados a aproximadamente 1% do tamanho da amostragem do sinal segundo (LETELLIER, 2008 *apud* ROCHA, 2013). Os demais operadores foram

arbitrados na metodologia por meio de testes preliminares que possibilitaram um melhor ajuste.

Abaixo, na Tabela 4, é mostrado o resultado do processamento do AG, de acordo com a configuração dos operadores genéticos da Tabela 3, com destaque para o resultado da Aptidão, que conforme o Capítulo 5, em relação ao Algoritmo Genético, representa a capacidade de cada indivíduo em solucionar o modelo proposto. O valor da maior Aptidão está acompanhado dos parâmetros utilizados na otimização que foram posteriormente adequados à simbolização das séries temporais.

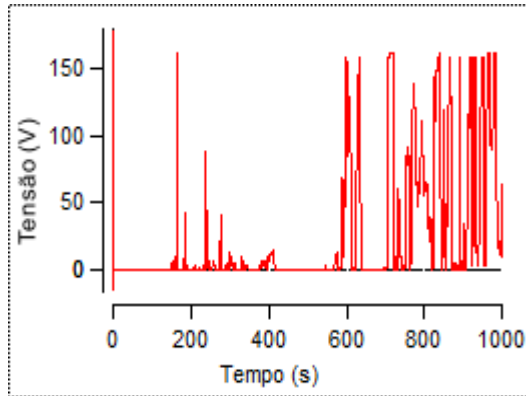
**Tabela 4 – Resumo da otimização da simbolização das séries temporais por Algoritmo genético.**

<i>DESCRIÇÃO</i>	<i>VALORES</i>
Aptidão Média	3,555
Menor Aptidão	1,160
Alfabeto - Menor Aptidão	4
Palavra - Menor Aptidão	4
Maior Aptidão	8,585
Alfabeto - Maior Aptidão	2
Palavra - Maior Aptidão	3
Tempo de Processamento	27h 39m 15s

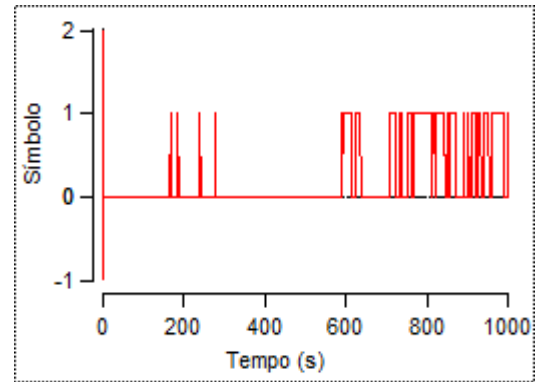
Verifica-se que nos valores apresentados na Tabela 4 a informação do indivíduo com maior aptidão entre as 1000 gerações, obteve aptidão igual a 8,585 entre os 100 mil indivíduos de todas as gerações, em que a aptidão média foi de 3,555 e a menor aptidão de 1,160.

## **8.2 ANÁLISE DA PPM POR DINÂMICA SIMBÓLICA**

O indivíduo com maior aptidão está configurado com os parâmetros *Tamanho do Alfabeto* ( $n=2$ ) e *Tamanho da Palavra* ( $m=3$ ) (Tabela 4). Esses valores determinaram a aplicação da transformação simbólica estática (conforme Equação 3 e Equação 4). O resultado da transformação simbólica sobre o sinal temporal pode ser observado na Figura 30, representando o eletrodo #5 do sinal de PPM do sujeito do GAU #1, e na Figura 31, do eletrodo #5 do sinal de PPM do sujeito #1 do GCO.

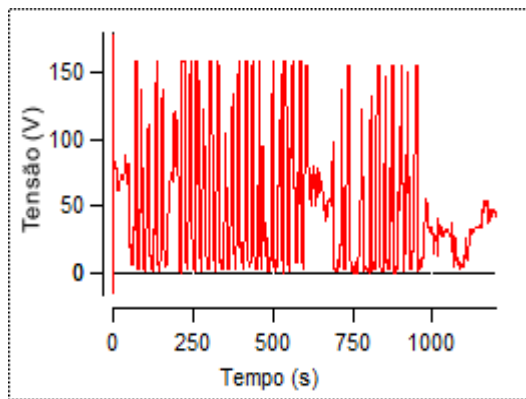


(a)

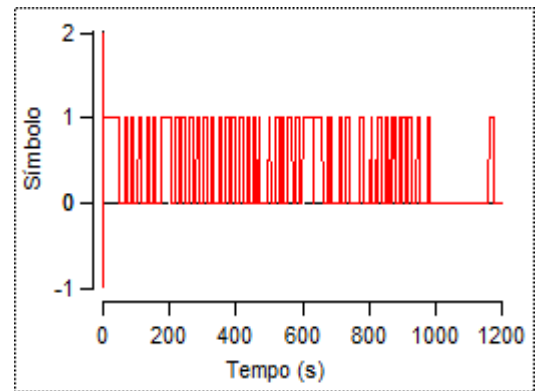


(b)

**Figura 30 – (a) Sinal do eletrodo #5 do sujeito do GAU #1 e (b) respectiva transformação simbólica.**



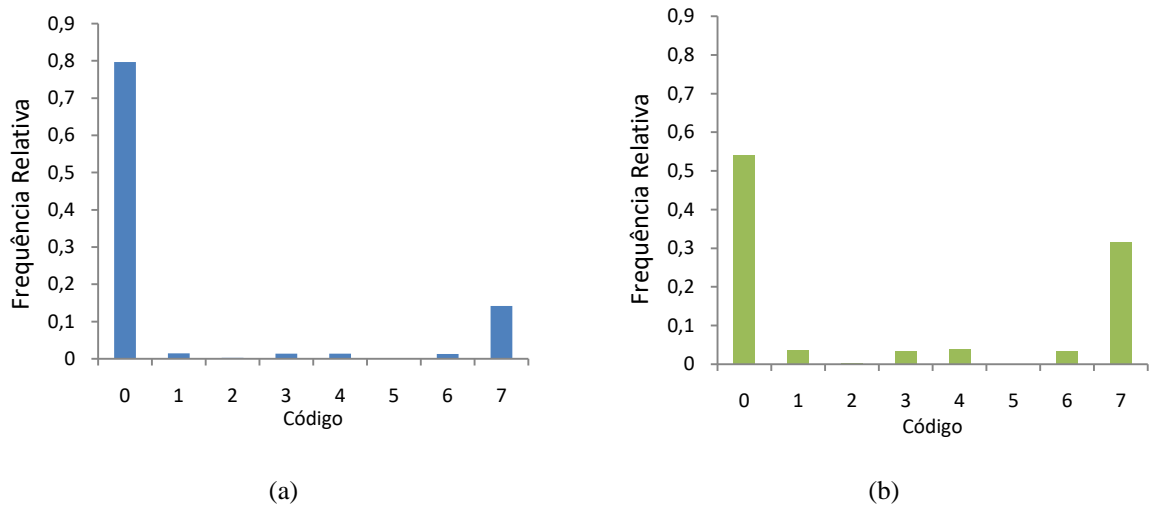
(a)



(b)

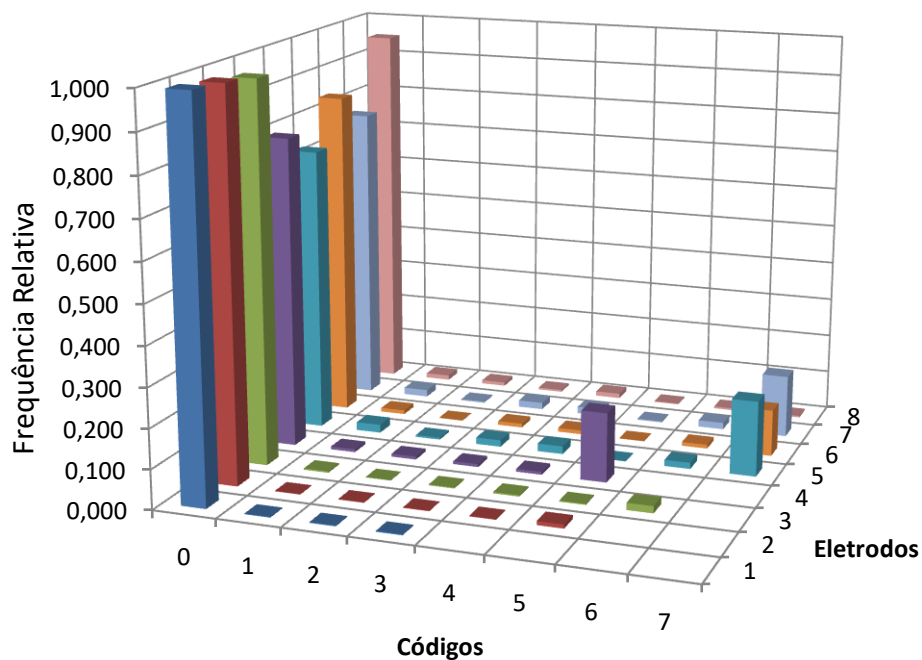
**Figura 31 – (a) Sinal do eletrodo #5 do sujeito #1 do GCO e (b) respectiva transformação simbólica.**

Após a transformação simbólica dos sinais, foi realizado o agrupamento das séries simbólicas, de três em três símbolos consecutivos (*palavras*), determinado pelos parâmetros do indivíduo com maior aptidão durante a otimização, conforme a Tabela 4. Cada grupo de símbolos binários foi transformado em códigos decimais e suas distribuições representadas por histogramas, como ilustrado nos exemplos da Figura 32.



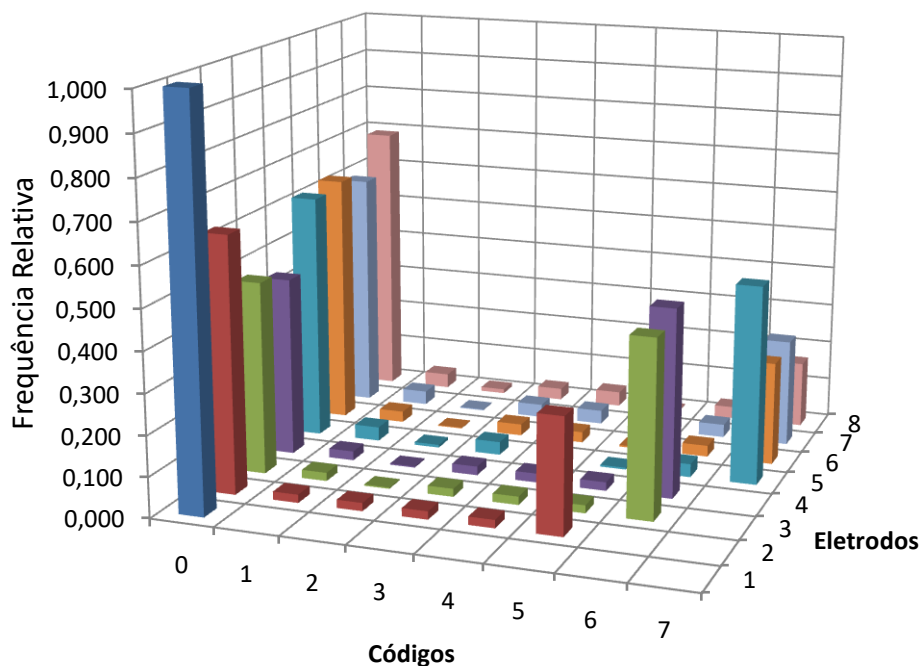
**Figura 32 – (a) Histograma da série simbólica do eletrodo #7 do sujeito GAU #1. (b) Histograma da série simbólica do eletrodo #7 do sujeito #3 do GCO.**

A Figura 33 ilustra os histogramas com o resultado das médias das frequências relativas dos sinais de PPM do GAU referentes às séries simbólicas de cada eletrodo.



**Figura 33 – Distribuição das médias das frequências relativas das séries simbólicas do GAU.**

A Figura 34 ilustra os histogramas com o resultado das médias das frequências relativas dos sinais de PPM do GCO referentes às séries simbólicas de cada eletrodo.



**Figura 34 – Distribuição das médias das frequências relativas das séries simbólicas do GCO.**

A partir destes histogramas é possível realizar diversos tipos de análises desta dinâmica do sistema. Percebe-se, na Figura 33 e Figura 34, que a probabilidade de ocorrência das frequências dos símbolos é maior para o GCO. Quanto maior a diferença de valores da distribuição entre os sinais entre os grupos, melhor a descrição do sistema obtido e maior a facilidade em se distinguir os grupos envolvidos.

A Figura 35 ilustra a média das frequências relativas das séries simbólicas entre todos os oito eletrodos para os dois tipos de sinais (GAU e GCO):

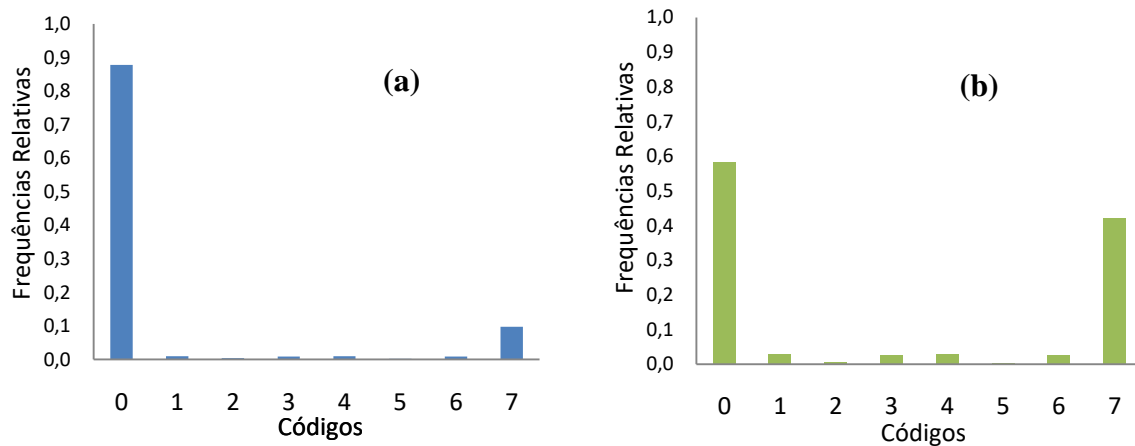


Figura 35 – Distribuição das médias das frequências relativas das séries simbólicas entre todos os eletrodos do GAU (a) e GCO (b).

### 8.3 ANÁLISE DA DINÂMICA SIMBÓLICA POR MEIO DE TESTES ESTATÍSTICOS

De acordo com os parâmetros utilizados na simbolização, os resultados das sequências e sequências e posteriormente os códigos delas derivados foram analisados estatisticamente para validar para validar a metodologia. As análises das porcentagens médias por símbolo nas amostras das séries das séries simbólicas estão descritas na Tabela 5 para séries do GAU, e

Tabela 6 para as séries do GCO. Os valores discriminam a frequência média de ocorrência dos símbolos em cada eletrodo, possibilitando, pelos dados, analisar quais símbolos possuem maior, menor ou mesmo nível de prevalência em determinado eletrodo.

Tabela 5 – Frequência média por símbolo em dados das séries simbólicas do GAU.

Eletrodo	SÍMBOLO							
	0	1	2	3	4	5	6	7
1	97,0%	0,5%	0,2%	0,3%	0,5%	0,0%	0,3%	1,2%
2	85,2%	1,5%	0,3%	1,3%	1,4%	0,2%	1,3%	8,9%
3	83,7%	1,1%	0,1%	1,0%	1,1%	0,0%	1,0%	12,1%
4	67,8%	1,8%	0,1%	1,8%	1,8%	0,1%	1,8%	24,9%
5	74,6%	2,1%	0,5%	1,8%	2,0%	0,2%	1,7%	17,2%
6	74,1%	1,9%	0,3%	1,8%	1,8%	0,2%	1,8%	18,2%
7	67,8%	2,2%	0,2%	2,2%	2,2%	0,2%	2,2%	23,1%
8	81,9%	1,7%	0,4%	1,5%	1,7%	0,2%	1,4%	11,2%

**Tabela 6 – Frequência média por símbolo em dados das séries simbólicas do GCO.**

<b>Eletrodo</b>	<b>SÍMBOLO</b>							
	<b>0</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>
1	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	0,0%	100,0%
2	62,1%	1,9%	0,0%	1,9%	1,9%	0,0%	1,9%	30,3%
3	46,9%	1,9%	0,0%	1,9%	1,9%	0,0%	1,9%	45,4%
4	47,6%	2,0%	0,1%	1,9%	2,0%	0,0%	1,9%	44,6%
5	40,3%	2,7%	0,5%	2,6%	2,8%	0,3%	2,7%	48,1%
6	63,6%	2,9%	0,3%	3,0%	2,9%	0,3%	2,9%	24,2%
7	59,3%	3,6%	0,2%	3,5%	3,5%	0,1%	3,4%	26,4%
8	64,8%	3,9%	0,9%	3,4%	4,0%	0,4%	3,4%	19,4%

A estatística descritiva é a etapa inicial da análise utilizada para descrever e resumir os dados. Nas Tabela 7 e Tabela 8 estão expressas as medidas descritivas das séries simbólicas do GAU e séries simbólicas do GCO respectivamente.

**Tabela 7 - Medidas descritivas do sinal por eletrodo das séries do GAU.**

<b>ELETRODOS</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
Média	0.0	0.2	0.6	1.2	24.7	10.9	12.0	0.2
Desv. Padrão	0.1	1.8	5.7	3.1	47.2	32.4	27.7	2.6
Mínimo	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1° quartil	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mediana	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
3° quartil	0.0	0.0	0.0	1.0	16.0	0.0	6.0	0.0
Máximo	3.0	17.0	66.0	24.0	162.0	161.0	161.0	77.0
Intervalo Confiança Inferior	0.0	0.1	0.3	1.1	22.7	9.5	10.8	0.1
Intervalo Confiança Superior	0.0	0.3	0.8	1.4	26.8	12.4	13.2	0.3

**Tabela 8 - Medidas descritivas do sinal por eletrodo das séries do GCO.**

<b>ELETRODOS</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>
Média	0	16.9	108.2	37.9	30.1	18.3	29.5	70.2
Desv. Padrão	0	22.7	90.6	28.7	39.9	31.3	39.4	73.7
Mínimo	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
1° quartil	0	0.0	9.0	8.0	5.0	0.0	4.0	0.0
Mediana	0	7.0	124.0	41.0	13.0	1.0	11.0	24.0
3° quartil	0	25.0	211.0	61.0	36.0	27.0	37.0	156.0
Máximo	0	97.0	219.0	102.0	159.0	159.0	158.0	159.0
Intervalo Confiança Inferior	0	16.2	105.1	37.0	28.7	17.3	28.2	67.7
Intervalo Confiança Superior	0	17.7	111.2	38.9	31.4	19.4	30.9	72.7

Com a configuração dos parâmetros da DS pelos valores otimizados pelo AG, observa-se que os histogramas dos indivíduos do GCO indicam frequências de ocorrência não nulas para um número maior de palavras, sugerindo que os indivíduos do GCO

apresentam um comportamento mais complexo do que os indivíduos do GAU. Isso pode ser observado na Tabela 7 e Tabela 8, nas quais se observa que a mediana dos sinais do GCO é maior do que no GAU, confirmando essa conjectura.

Na Figura 36 é apresentado o gráfico *boxplot* com as séries simbólicas dos grupos sem a subdivisão por eletrodo. *Boxplot* é utilizado para uma comparação visual entre dois ou mais grupos e, conforme (FARIAS, 2008), permitindo avaliar a distribuição empírica de dados. Neste gráfico duas ou mais caixas são colocadas lado a lado e se compara a variabilidade entre elas, a mediana e assim por diante. Outro ponto importante é a diferença entre os quartis, que é uma medida da variabilidade dos dados. No *boxplot* ainda são observados, da esquerda à direita: o mínimo aceitável, 1º quartil, mediana (2º quartil), 3º quartil e o máximo aceitável para essa distribuição, podendo ainda ser observados os pontos discrepantes (*outliers*). De acordo com o formato do gráfico da Figura 36, percebe-se que existe grande diferença entre as séries simbólicas que cada grupo representa, pois os sinais estão distribuídos de forma muito diferente no GCO em relação aos GAU, devido à concentração dos dados do GAU serem maiores se observada a ausência da caixa do *boxplot* em sua representação.

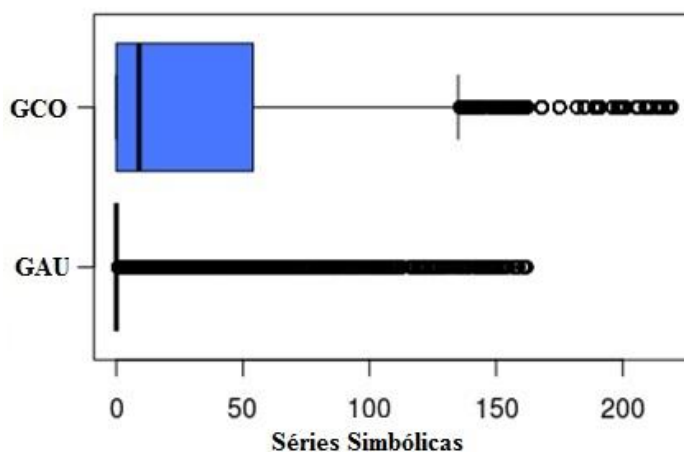


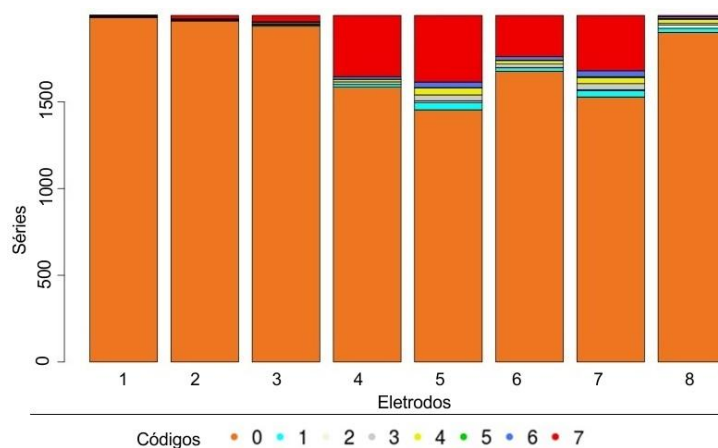
Figura 36 – Gráfico *Boxplot* da distribuição dos sinais por grupo.

Analisando o gráfico *boxplot* (Figura 36) pode-se observar que os dados das séries simbólicas do grupo GAU possuem menor variabilidade em relação ao grupo GCO. Pode-se considerar que os dados das séries simbólicas entre os grupos GCO e GAU possuem alto grau de variabilidade entre ambos, propondo que a comparação entre as séries simbólicas dos grupos possuem variâncias significativamente diferentes. Essa observação foi realizada pelos valores da Tabela 7 e Tabela 8 comprovada visualmente pelo gráfico e

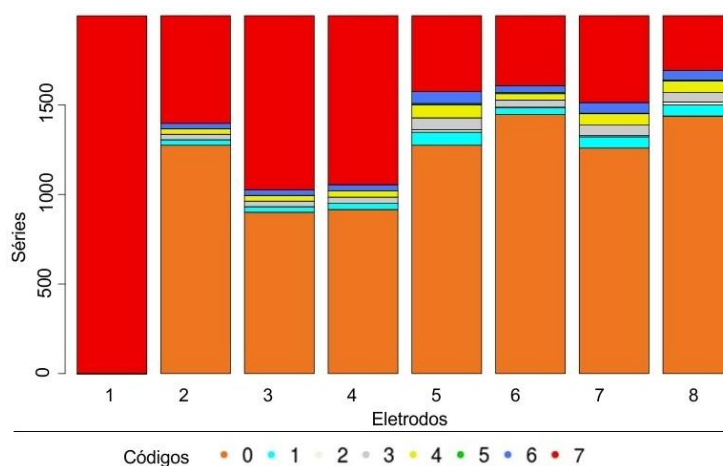


confirmada pelo Teste de *Qui-quadrado* (Tabela 9), em que os valores de referência foram menores que os valores de significância (*p-valores*), comprovando a diferença entre as séries.

A Figura 37 e Figura 38 ilustram a distribuição dos símbolos em cada eletrodo com suas respectivas frequências, possibilitando a análise visual do nível de ocorrência dos símbolos agrupados para cada eletrodo.



**Figura 37 - Distribuição dos símbolos nos eletrodos das séries do GAU.**



**Figura 38 - Distribuição dos símbolos nos eletrodos das séries do GCO.**

Para as séries do grupo GAU, conforme Figura 37, o símbolo “0” possui um maior grau de prevalência entre todos os oito eletrodos, seguido do símbolo “7” com destaque nos eletrodos 4 a 7. Para as séries do grupo GCO, conforme Figura 38, o símbolo “7” possui destaque em todos os eletrodos, seguido da incidência significativa do símbolo “0”

nos eletrodos 2 a 8. Observa-se nas figuras citadas que os símbolos 0 a 7 são melhor distribuídos entre os eletrodos do grupo GCO em comparação com o grupo GAU.

O teste de *Qui-quadrado* serve para avaliar quantitativamente a relação entre o resultado de um experimento e a distribuição esperada para o fenômeno, isto é, ele nos diz com quanta certeza os valores observados podem ser aceitos como os regidos pela teoria em questão. A Tabela 9 mostra o teste comparando os símbolos pelos grupos e pelos eletrodos e como os valores dos *p-valor* foram abaixo de 0.0001, sendo arredondados para 0, pode-se então concluir que existe diferença significativa entre as séries GAU e GCO para os 8 eletrodos.

**Tabela 9 - Teste Qui-quadrado.**

	<i>Qui-quad.</i>	<i>g.l.</i>	<i>p-valor</i>
Grupos	5453065	7	0
Eletrodos	5954116	49	0

Comparou-se para cada eletrodo o Teste *t Student* avaliando se existia diferença das médias entre os grupos, conforme apresentado na Tabela 10 a Tabela 17.

Foi feita uma avaliação das séries simbólicas com o teste *t Student* para todas as amostras dos dois grupos (GAU e GCO). Observou-se a ocorrência de uma diferença significativa (com 95% de confiança) entre as médias, das séries simbólicas dos códigos de “0” a “7”, para os dois grupos analisados. Essa diferença significativa foi identificada nos eletrodos 1, 2, 3, 4 e 7. No eletrodo 5, os códigos 4, 6 e 7 não apresentaram diferença significativa que permitissem distinguir sinais dos grupos GAU e GCO. Da mesma forma, no eletrodo 6 o código 7 também não apresentou diferença significativa que permitisse distinguir sinais dos dois grupos.

**Tabela 10 – Comparação estatística *t* Student das frequências de símbolos do eletrodo #1.**

<b><i>Eletrodo #1</i></b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0	0	-	-	-	-
001	1	0	0	-	-	-	-
010	2	0	0	-	-	-	-
011	3	-	-	-	-	-	-
100	4	2	1,1547	-	-	-	-
101	5	-	-	-	-	-	-
110	6	-	-	-	-	-	-
111	7	-	-	-	-	-	-

**Tabela 11 - Comparação estatística *t* Student das frequências de símbolos do eletrodo #2.**

<b><i>Eletrodo #2</i></b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0	0	3,42	4,74	-33,57	≤ 0,05
001	1	0	0	2,384	3,911	-4,39	≤ 0,05
010	2	-	-	-	-	-	-
011	3	0	0	2,384	3,911	-4,39	≤ 0,05
100	4	16	1,732	53,98	20,507	-12,59	≤ 0,05
101	5	-	-	-	-	-	-
110	6	16	1,732	53,98	20,507	-12,59	≤ 0,05
111	7	15,35	1,531	43,3	19,59	-39,8	≤ 0,05

**Tabela 12 - Comparação estatística *t* Student das frequências de símbolos do eletrodo #3.**

<b><i>Eletrodo #3</i></b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0	0	21,402	30,562	-27,39	≤ 0,05
001	1	0	0	17,792	30,764	-4,21	≤ 0,05
010	2	0	0	-	-	-	-
011	3	0	0	17,792	30,764	-4,21	≤ 0,05
100	4	18,857	25,673	181,67	40,875	-14,55	≤ 0,05
101	5	-	-	-	-	-	-
110	6	20	28,556	181,67	40,875	-11,6	≤ 0,05
111	7	24,333	30,617	189,6	36,633	-31,89	≤ 0,05

Tabela 13 - Comparação estatística *t Student* das frequências de símbolos do eletrodo #4.

<b>Eletrodo #4</b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0,11	0,358	12,576	13,308	-36,93	≤ 0,05
001	1	0,733	0,883	12,237	11,084	-7,87	≤ 0,05
010	2	-	-	-	-	-	-
011	3	0,866	0,915	12,237	11,084	-7,77	≤ 0,05
100	4	5,2	3,144	67,305	16,267	-27,14	≤ 0,05
101	5	-	-	-	-	-	-
110	6	5,866	3,09	67,305	16,267	-27,14	≤ 0,05
111	7	6,039	4,621	62,243	14,668	-127,4	≤ 0,05

Tabela 14 - Comparação estatística *t Student* das frequências de símbolos do eletrodo #5.

<b>Eletrodo #5</b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	1,618	3,82	10,008	9,54	-36,8	≤ 0,05
001	1	2,761	4,514	12,895	12,31	-7,75	≤ 0,05
010	2	4	8,432	16,13	8,319	-3,81	≤ 0,05
011	3	10,411	8,442	20,504	15,144	-4,96	≤ 0,05
100	4	69,15	42,24	60,056	39,986	1,19	≥ 0,05*
101	5	26	0	66,846	45,441	-3,24	≤ 0,06
110	6	102,44	48,538	87,778	43,235	1,54	≥ 0,05*
111	7	104,81	43,676	80,423	45,413	8,71	≥ 0,05*

Tabela 15 - Comparação estatística *t Student* das frequências de símbolos do eletrodo #6.

<b>Eletrodo #6</b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0,057	0,567	4,32	8,292	-25,4	≤ 0,05
001	1	1,3	2,754	9,65	10,284	-5,82	≤ 0,05
010	2	0	0	13	14,764	-2,32	≤ 0,05
011	3	1,4	2,21	17,238	10,475	-11,5	≤ 0,05
100	4	46,8	41,55	57,803	38,023	-1,04	≤ 0,05
101	5	27	0	32,181	4,622	-3,71	≤ 0,05
110	6	73,9	46,535	87,169	54,849	-1,06	≤ 0,05
111	7	81,067	48,091	60,325	34,503	6,11	≥ 0,05*

**Tabela 16 - Comparação estatística *t Student* das frequências de símbolos do eletrodo #7.**

<b>Eletrodo #7</b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0,966	2,262	8,873	9,982	-35,4	≤ 0,05
001	1	3	3,978	16,084	14,897	-8,25	≤ 0,05
010	2	4,333	5	24,636	16,268	-3,82	≤ 0,05
011	3	5,171	5,42	21,821	14,468	-9,75	≤ 0,05
100	4	27	19,411	72,476	47,587	-8,04	≤ 0,05
101	5	20,2	3,898	21,6	2,19	-0,7	≤ 0,05
110	6	42,764	31,292	102,46	46,623	-8,41	≤ 0,05
111	7	61,523	38,361	71,54	41,691	-3,86	≤ 0,05

**Tabela 17 - Comparação estatística *t Student* das frequências de símbolos do eletrodo #8.**

<b>Eletrodo #8</b>							
Sequência	Símbolo	GAU		GCO		Estatística <i>t</i>	
		Média	Desvio Padrão	Média	Desvio Padrão	<i>t</i> calc	<i>t</i> 0,05 unilateral
000	0	0	0	68,705	77,025	-44,1	≤ 0,05
001	1	0	0	72,3	76,539	-9,72	≤ 0,05
010	2	0	0	144,85	40,194	-18,7	≤ 0,05
011	3	0	0	54,32	71,133	-7,2	≤ 0,05
100	4	2,75	3,2	87,311	66,156	-13,3	≤ 0,05
101	5	1	0	100,2	75,503	-4,15	≤ 0,05
110	6	19,8	30,56	83,195	59,051	-5,53	≤ 0,05
111	7	9	2,39	69,36	58,258	-22,4	≤ 0,05

A Tabela 18 apresenta o resultado das estatísticas quantitativas da distribuição das séries simbólicas de cada grupo, analisando a diferença entre os histogramas gerados pela Dinâmica Simbólica. Em seguida a Figura 39 ilustra os resultados descritos na Tabela 18, possibilitando visualizar a as diferentes análises estatísticas quantitativas na comparação dos histogramas. Essas análises, como a *Entropia de Shannon*, *Norma Euclidiana* e  $\chi^2$  *modificado*, são métodos quantitativos que permitem comparar diferenças entre os histogramas dos dois grupos estudados.

Tabela 18 - Estatística quantitativa da distribuição das séries simbólicas.

<i>Eletrodo</i>	<i>Freq. GAU</i>	<i>Freq. GCO</i>	<i>Entropia GAU</i>	<i>Entropia GCO</i>	<i>Norma Euclidiana</i>	<i>X<sup>2</sup> Modificado</i>
1	0,878	0,582	0,039386	0	1,409981573	2,000000018
2	0,009	0,028	0,061287	0,595998	0,46259977	0,402200964
3	0,003	0,006	0,094201	0,599132	0,685673081	0,646298295
4	0,009	0,026	0,356018	0,63117	0,438992981	0,232428146
5	0,009	0,028	0,450254	0,389297	0,094909308	0,029035676
6	0,002	0,004	0,307625	0,579164	0,098676227	0,022009815
7	0,009	0,026	0,416356	0,635917	0,161406989	0,045923647
8	0,097	0,421	0,145855	0,540848	0,291619974	0,235714028

Os dados da Tabela 18 ilustrados pela Figura 39, fornecem informações que confirmam estatisticamente, pelo distanciamento dos resultados, quando as sequências de símbolos podem efetivamente discriminar a diferença entre os grupos. Ressalte-se que apenas os oitos valores inteiros da abscissa do gráfico têm significado nas curvas apresentadas.

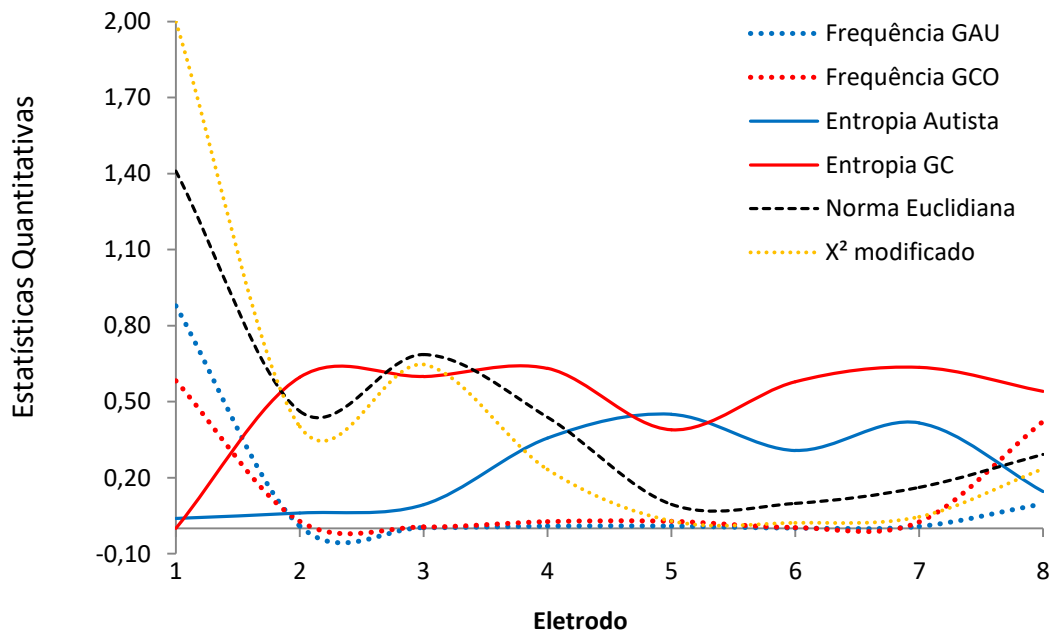


Figura 39 – Gráfico da Tabela 18 ilustrando os resultados das estatísticas quantitativas.

O teste *t Student* indicou estatisticamente que os eletrodos #5 e #6 possuem valores de frequências relativas que não diferenciam os dois tipos de sinais. Este resultado é confirmado pelos demais indicadores estatísticos quantitativos. Observando-se visualmente

a Figura 39 os valores da *Entropia de Shannon* para os dois tipos de sinais se aproximam, para o caso dos eletrodos #5 e #6, indicando uma baixa participação destes dois eletrodos na caracterização dos dois grupos estudados. Outra confirmação deste resultado está nos valores do resultado da *Norma Euclidiana* e do  *$x^2$  modificado*, onde não se verificou distinção clara nos testes estatísticos, quando comparados os dois grupos. Isso indica que para a distinção entre os grupos avaliados (GAU e GCO), há valores equiprováveis nas frequências relativas observadas nos eletrodos #5 e #6.

#### **8.4 CLASSIFICAÇÃO AUTOMÁTICA DAS SÉRIES SIMBÓLICAS POR REDE NEURAL ARTIFICIAL**

Após o término da simbolização dos sinais, foi construída, treinada e testada a RNA de forma a permitir a classificação automática de determinado sinal, como pertencendo ao GAU ou ao GCO.

Inicia-se o processo de classificação automática das séries simbólicas, realizando a interpolação quadrática dos dados de PPM. O método permite construir um novo conjunto de dados a partir de um conjunto discreto de dados pontuais conhecidos (SANTOS *et al.*, 2010; ARENALES *et al.*, 2007; SOARES, 2006), proporcionando um maior conjunto de dados para treinamento e testes da RNA. Este procedimento não é o mais adequado, mas permite aumentar o número de indivíduos do grupo testado buscando maior representatividade estatística. Futuramente, testes com grupos maiores deverão ser conduzidos para se confirmar os resultados.

Para construção da RNA, definiram-se as configurações conforme Tabela 19, baseando-se em testes preliminares para o ajustamento e redução da taxa de erro durante o treinamento (HAYKIN, 2001).

**Tabela 19 – Parâmetros para construção da RNA algoritmo MLP**

<b>CONSTRUÇÃO DA REDE MLP</b>	
Número de épocas	9000
Neurônios da camada oculta	3
Taxa de aprendizagem	90%
Taxa de inércia	30%
Arquivo de conhecimento	Perfil.txt

A Tabela 20 lista os dados utilizados para o treinamento da RNA. No treinamento a rede foi alimentada com as entropias de cada eletrodo e seu respectivo *Perfil* (1 – GAU e 2 - GCO), o perfil indica qual a saída desejada quando se informa os valores das entropias dos oito eletrodos à RNA.

**Tabela 20 - Conjunto de treinamento para RNA.**

<i>Sinal</i>	<i>Perfil</i>	<i>ELETRODO</i>							
		1	2	3	4	5	6	7	8
1	1	0,039	0,036	0,037	0,169	0,450	0,308	0,374	0,146
2	1	0,039	0,087	0,151	0,543	0,450	0,308	0,459	0,146
3	1	0,039	0,226	0,243	0,546	0,450	0,308	0,459	0,237
4	1	0,039	0,306	0,285	0,540	0,356	0,466	0,464	0,237
5	1	0,039	0,322	0,391	0,594	0,517	0,466	0,464	0,335
6	1	0,039	0,294	0,391	0,407	0,517	0,466	0,674	0,489
7	1	0,039	0,557	0,231	0,567	0,517	0,597	0,674	0,489
8	1	0,039	0,557	0,557	0,567	0,517	0,597	0,674	0,489
9	1	0,474	0,557	0,557	0,567	0,194	0,597	0,674	0,489
10	1	0,021	0,147	0,188	0,408	0,300	0,232	0,294	0,225
11	2	0	0,579	0,664	0,586	0,554	0,139	0,570	0,560
12	2	0	0,579	0,664	0,586	0,552	0,520	0,570	0,560
13	2	0	0,579	0,666	0,586	0,669	0,520	0,750	0,560
14	2	0	0,524	0,666	0,586	0,669	0,743	0,750	0,576
15	2	0	0,524	0,582	0,655	0,689	0,743	0,460	0,651
16	2	0	0,673	0,570	0,637	0	0,622	0,754	0,651
17	2	0	0,673	0,484	0,646	0	0,638	0,705	0,651
18	2	0	0,673	0,654	0,637	0	0,743	0,750	0,651
19	2	0	0,539	0,525	0,602	0,544	0,597	0,495	0,476
20	2	0	0,576	0,619	0,655	0,624	0,398	0,663	0,495

Com os dados da Tabela 20 o treinamento foi realizado atingindo um erro total de 0,000001. Em seguida foi realizada a sessão de testes, a rede foi carregada com o arquivo de conhecimento (Perfil.txt – conforme Tabela 19) salvo para efetuar as classificações conforme os padrões apresentados à sua entrada; assim os resultados da classificação durante a sessão de testes foram (ver Tabela 21):



**Tabela 21 – Resultado da classificação dos sinais pela RNA.**

<i>Resultado do teste da Rede Neural Artificial</i>										
Parâmetros da RNA				Taxa de Erro	Tamanho das amostras		Resultado do GAU		Resultado do GCO	
Épocas	Neurônios camada oculta	Taxa aprendizagem	Taxa inércia	Erro Total	GAU	GCO	Acerto	Erro	Acerto	Erro
<b>9000</b>	<b>3</b>	<b>90%</b>	<b>30%</b>	<b>0,000001</b>	<b>10</b>	<b>10</b>	<b>10</b>	<b>0</b>	<b>8</b>	<b>2</b>

## 8.5 ANÁLISE ESTATÍSTICA DA CLASSIFICAÇÃO POR RNA

Os resultados referentes à fase de testes da RNA (ver Tabela 21) foram analisados estatisticamente quanto ao seu potencial de identificar de forma automática os padrões das séries simbólicas. E ficou constatado que é possível realizar a distinção dos comportamentos dinâmicos do GAU comparado ao GCO por meio da simbolização das séries temporais. Tendo sido utilizadas, para a fase treinamento, os parâmetros listados na Tabela 19. A capacidade da RNA em classificar novos casos nos dois grupos de dados foi testada e foi obtida uma taxa de precisão de acertos de 91,66%, como apresentado na Tabela 22.

**Tabela 22 – Resultados quantificados dos testes da RNA.**

<i>REDE NEURAL MLP</i>	
DESCRIÇÃO	VALORES
Acertos	90%
Erros	10%
Verdadeiros positivos	100%
Falsos negativos	0%
Verdadeiros negativos	80%
Falsos positivos	20%
Especificidade	80%
Sensibilidade	100%
Seletividade positiva	83,33%
Seletividade negativa	100%
<b>Precisão</b>	<b>91,66%</b>

## 9 – CONCLUSÃO

---

Uma metodologia baseada em técnicas de reconhecimento de padrões foi apresentada, com o intuito de avaliar um possível biomarcador que possa auxiliar no diagnóstico precoce do autismo através da investigação da PPM, com base em uma conjuntura simbólica dinâmica. A principal descoberta é que a abordagem dinâmica, utilizando o método de otimização por AG aplicada aos parâmetros da simbolização das séries temporais, revela diferenças significativas entre as amostras de PPM.

O trabalho traz a análise de uma proposta inédita, uma vez que não há na literatura comentários sobre a otimização dos parâmetros da DS em dados de PPM para validar um potencial candidato a biomarcador para o autismo.

Para isso, a estruturação dos dados em uma base de dados confiável é a etapa inicial, primordial para a metodologia empregada. Através da estruturação dos dados foi possível confiar os dados de PPM ao processamento exaustivo exigido durante a otimização pelo AG, na transformação simbólica e no treinamento/teste executado pela RNA. O fluxo de dados executado na entrada e saída das informações são relativamente grandes e qualquer erro de estruturação e processamento causaria falhas nos resultados ou perda de dados, podendo condenar a proposta metodológica caso haja falhas nesta etapa. A aplicação do Sistema Gerenciador de Banco de Dados PostGreSQL na tarefa de estruturar os dados, demonstrou confiabilidade e facilidade na manipulação, além de permitir futuras atualizações do método.

A Dinâmica Simbólica é um método capaz de analisar sinais periódicos e caóticos, porém há, em sua aplicação, uma dificuldade na configuração correta de seus parâmetros. Devido à regra de particionamento ser equiprovável, as frequências relativas de cada

possível sequência de símbolos para dados verdadeiramente aleatórios tenderá a ser igual; assim, qualquer desvio significativo a partir de sequências equiprováveis é indicativo de correlação de tempo e estrutura determinística. Foi proposto, por meio desta discussão, um modelo que seja suficiente para determinar as melhores combinações de parâmetros para a DS, solucionando essa lacuna do método. Para determinar os valores destes parâmetros utiliza-se um modelo de otimização por Algoritmo Genético, tratando de forma otimizada esta combinação, em que os resultados possam ser precisos e com qualidade estatística confiável.

Com a implementação da otimização no processo de simbolização das séries temporais, a variabilidade experimental – isto é, as variações dos parâmetros da DS para comparação de diversos histogramas – é descartada, não necessitando da definição de limites de confiança para concretizar os métodos estatísticos aplicados na simbolização.

Analisando o resultado da otimização, percebe-se que há variação entre a menor e maior aptidão durante o processamento de otimização, – em que são testadas e validadas inúmeras combinações –, confirmando a hipótese de que as variações dos parâmetros da dinâmica simbólica influenciam no resultado das amostras. Essa combinação é importante para que o resultado da simbolização possa refletir as estruturas determinísticas propostas pelo método. Caso uma configuração escolhida não seja a ideal para amostragem, a análise estatística dos dados pode não determinar diferenças significativas em suas análises.

Com a análise dos histogramas resultantes da simbolização otimizada por AG, os valores estatísticos comprovaram o distanciamento entre as séries simbólicas, e esses resultados inspiraram a implementação de um método de RNA de classificação automática para determinar, em um sinal, as características de um determinado grupo. O método de classificação conseguiu diferenciar as amostras com uma precisão de 91,66%.

O método inova em analisar os dados de PPM de GAU, demonstrando através dos resultados que há padrões de PPM de autistas, podendo auxiliar no diagnóstico do autismo como um biomarcador importante.

Diante dos resultados iniciais obtidos, conclui-se que o método de simbolização dinâmica de séries temporais otimizadas por AG foi capaz de detectar diferenças estatisticamente significativas das séries temporais de PPM GAU e GCO. O método se

mostrou promissor, no entanto, uma vez que os resultados obtidos consideraram uma pequena amostra de sinais, o estudo deveria ser ampliado em futuras pesquisas, de forma a considerar um grupo maior de amostras e validar os resultados obtidos. Um dedicado esforço ainda é necessário para transformar esse método de pesquisa em um biomarcador clinicamente capaz de determinar o risco de autismo antes do início dos sintomas mais severos.

## 10 – PERSPECTIVAS

---

Tem-se perspectivas de desenvolvimento de trabalhos futuros decorrentes desta pesquisa de doutorado. Uma delas é a análise dos eletrodos, individualmente na PPM, verificando aqueles cujo posicionamento melhor contribuem para a identificação das características do GAU e GCO.

Também poderá ser ampliado o método com a análise em sensores acoplados em todos os membros inferiores e superiores, além da análise postural durante a marcha. Assim, a metodologia poderá concluir com maiores detalhes os défices da coordenação motora dos indivíduos.

Acompanhando a linha de pesquisa que envolve o autismo, aplicar o método na análise dos sinais de Eletroencefalograma para analisar se há diferentes características entre os sinais de pessoas que possuem o autismo. Visto que há pesquisas que propõem estudos destes sinais e o método pode auxiliar no avanço destas pesquisas.

A metodologia também poderá auxiliar na análise de PPM de atletas acometidos de lesões do joelho. Estatísticas apontam que 40% dos atletas que sofrem lesões em um dos joelhos, ao retornarem suas atividades sofrem lesões no outro joelho. A possibilidade da análise da PPM destes atletas durante sua recuperação, com o auxílio das palmilhas para investigar se há algum fator de controle postural ou muscular durante sua recuperação poderá ajudar a descobrir algum desequilíbrio postural ou muscular causado pela lesão.

# 11 – REFERÊNCIAS

---

1. ABRANTES, J. M. C. S. & L. F. F. SANTOS. *Plantar pressure assessment: A new tool for postural instability diagnosis in Multiple Sclerosis*. Technologies for Medical Sciences. 2011.
2. ALAJMI, A. WRIGHT, J. *Selecting the Most Efficient Genetic Algorithm Sets in Solving Unconstrained Building Optimization Problem*. International Journal of Sustainable Built Environment, v. 3, n. 1, p. 18–26, 2014.
3. ANOCHI, J. A. *Previsão climática de precipitação por redes neurais autoconfiguradas*. Tese de doutorado em computação aplicada. Instituto Nacional de Pesquisas Espaciais – INPE. São José dos Campos. 2015.
4. ARENALES, S. DAREZZO, A. *Cálculo numérico: Aprendizagem com apoio de software*. São Paulo: Cengage Learning. 2007.
5. ARRUDA, G. P.; DEMATTE, J. A. M.; CHAGAS, C. S. *Mapeamento digital de solos por redes neurais artificiais com base na relação solo-paisagem*. Revista Brasileira de Ciências do Solo, v.37, p.327-338, 2013
6. ASTRO M. P, SOARES D, MENDES E, MACHADO L. *The influence of different in-shoe inserts on the plantar pressure during the gait of healthy elderly people*. Gait Posture. 2012.
7. BANDEIRA, A. F. *Aplicação de rede neural artificial para o reconhecimento do diabetes mellitus gestacional com marcadores não-glicêmicos*. Dissertação de Mestrado em Bioinformática. Universidade Federal do Paraná. 2015.
8. BAUMERT M, JAVORKA M, KABIR M. 2015. *Joint symbolic dynamics for the assessment of cardiovascular and cardiorespiratory interactions*. Phil. Trans. R. Soc. A 373: 20140097.
9. BECCENERI, J. *Meta-heurística e otimização combinatória: aplicações em problemas ambientais*. São José dos Campos, 2013. 31–42 p. 21p. Mini Curso de Escola de Verão.
10. BEGEER, S. M, DAVID. W, Bernadette. VENDERBOSCH, S. R, DORIEN. STEKELENBURG, F. HANS M. *Sex Differences in the Timing of Identification Among Children and Adults with Autism Spectrum Disorders*. Journal Autism Dev Disord (2013) 43:1151–1156 DOI 10.1007/s10803-012-1656-z

11. BHAT, A. N. LANDA, R. J. & GALLOWAY, J. C. *Current Perspectives on Motor Functioning in Infants, Children, and Adults With Autism Spectrum Disorders*. Physical Therapy, 91 (7), 1116-1129. 2011.
12. BING, L. CHEN, D. YANG, Y. YU, G. *Effect of second toe-to-hand transfer on the plantar pressure distribution of the donor foot*. Acta ortop. bras. vol.24 no.1 São Paulo Jan./Feb. 2016
13. BOLTON, P. F, GOLDING, J. EMOND, A. STEER, C. D. *Autism spectrum disorder and autistic traits in the Avon Longitudinal Study of Parents and Children: precursors and early signs*. J Am Acad Child Adolesc Psychiatry. 2012. 51(3):249–260.e25.
14. BRAS, G. P. *Estudo do Perfil Motor de Crianças com Perturbações do Espectro do Autismo*. Dissertação de mestrado em Ciência do Desporto. Faculdade do Desporto. Universidade de Porto. Porto – Portugal. 2009.
15. CANDAN, G. YAZGAN, H. R. *Genetic algorithm parameter optimisation using Taguchi method for a flexible manufacturing system scheduling problem*. International Journal of Production Research, v. 53, n. 3, p. 897–915, 2014.
16. CASTRO, M. P., (2010) - *Análise das forças e pressões plantares durante a marcha de pessoas com amputação transfemoral*, MSc. Faculdade de desporto da Universidade do Porto, Porto.
17. CHANG, C. H. WADE, M. G. STOFFREGEN, T. A. HSU, C. Y. PAN, C. Y. *Visual tasks and postural sway in children with and without autism spectrum disorders*. Research in Developmental Disabilities, 31, 1536-1542. 2010.
18. CHIU M. C, WU H. C, CHANG L. Y, WU M. H. *Center of pressure progression characteristics under the plantar region for elderly adults*. Gait Posture. 2013.
19. CLIFFORD SM, HUDRY K, ELSABBAGH M, CHARMAN T, JOHNSON MH; BASIS Team. *Temperament in the first 2 years of life in infants at high-risk for autism spectrum disorders*. J Autism Dev Disord. 2013;43(3):673–686
20. COSTA, A. A. *Investigação de modelos comportamentais de ratos por meio de algoritmos genéticos*. Tese de doutorado. Universidade de São Paulo. Ribeirão Preto – São Paulo. 2015.
21. COSTA, D. RIBEIRO, J. ROSA, S. *Análise de temperatura usando dinâmica simbólica na prevenção de câncer esofágico*. XXIV Congresso Brasileiro de Engenharia Biomédica. 2014.
22. CROWTHER, R. FERNANDO, M. LAZZARINI, P. SANGLA, K, CUNNINGHAM, M, BUTTNER, P. *Biomechanical characteristics of peripheral diabetic neuropathy: A systematic review and meta-analysis of findings from the gait cycle, muscle activity and dynamic barefoot plantar pressure*. Clinical Biomechanics. 2013.
23. DANIELS, A. M. MANDELL, D. S. *Explaining differences in age at autism spectrum disorder diagnosis: a critical review*. Autism. 2014;18 (5):583–597
24. DAW, C.S., FINNEY, C.E. E TRACY, E.R. 2003. *A review of symbolic analysis of experimental data*, Review of Scientifics Instruments, vol. 74, pp. 916-930.

25. DAW, C.S., KENNEL, M.B., FINNEY, C.E. E CONNOLLY, F.T., 1998, “*Observing and modeling nonlinear dynamics in an internal combustion engine*”, Physical Review E, vol. 57,n. 3, pp.2811-2819.
26. DAW. C.S. FINNEY. C.E.A. *A review of symbolic analysis of experimental data*. 2002.
27. DEEPASHINI H, OMAR B, PAUNGMALI A, AMARAMALAR N, OHNMAR H, LEONARD J. *An insight into the plantar pressure distribution of the foot in clinical practice: narrative review*. Polish Ann Medicine. 2014;21:51-6.
28. DEIRDRE, K. LOUISE, G. JOHN, G. *Gait Deviations in Children with Autism Spectrum Disorders: A Review*. Autism Research and Treatment. 2015. Article ID 741480, 8 pages <http://dx.doi.org/10.1155/2015/741480>.
29. ELSABBAGH M, FERNANDES J, JANE Webb S, DAWSON G, CHARMAN T, JOHNSON MH; British Autism Study of Infant Siblings Team. *Disengagement of visual attention in infancy is associated with emerging autism in toddlerhood*. Biol Psychiatry. 2013;74(3):189– 194
30. ESPOSITO, G. VENUTI, P. *Analysis of toddler’s gait after six months of independent walking to identify autism: a preliminary study*. Perceptual and Motor Skills. 2008. 106:259–69.
31. ESPOSITO, G. VENUTI, P. APICELLA, F. MURATORI, F. *Analysis of unsupported gait in toddlers with autism*. Brain and Development 2011. 33:367–73.
32. ESPOSITO, G. VENUTI, P. MAESTRO, S. MURATORI, F. *An exploration of symmetry in early autism spectrum disorders: analysis of lying*. Brain and Development. 2009. 31:131–8.
33. ESPOSITO, G., and VENUTI, P. (2008). Analysis of toddlers’ gait after six months of independent walking to identify autism: a preliminary study. Percept. Mot. Skills 106, 259–269. doi:10.2466/pms.106. 1.259-269.
34. FARIAS, A. M. L.; Laurencel, L. C. *Estatística Descritiva*. Departamento de Estatística. Niterói-RJ. Universidade Federal Fluminense. 2008.
35. FEIN, D. BARTON, M. EIGSTI, I.M. *Optimal outcome in individuals with a history of autism*. J Child Psychol Psychiatry. 2013. 54: 195-205.
36. FERNANDO, M. CROWTHER R. G. CUNNINGHAM M. LAZZARINI P. A, SANGLA K.S. BUTTNER, P. GOLLEDGE, J. *The reproducibility of acquiring three dimensional gait and plantar pressure data using established protocols in participants with and without type 2 diabetes and foot ulcers*. Journal of Foot and Ankle Research. 2016. 9:4 DOI 10.1186/s13047-016-0135-8
37. FLANAGAN, J. E. LANDA, R. BHAT, A. BAUMAN, M. *Head lag in infants at risk for autism: a preliminary study*. American Journal of Occupational Therapy 2012. 66:577–85.
38. FLORES, J. H. F. *Comparação de modelos mlp/rna e modelos box-jenkins em séries temporais não lineares*. Dissertação de mestrado em Engenharia de Produção da Universidade Federal do Rio Grande do Sul. 2009.
39. GIACOMOZZI C, KEIJSERS N, PATAKY T, ROSENBAUM D. *International scientific consensus on medical plantar pressure measurement devices: technical requirements and performance*. Annali dell’Istituto superiore di sanita. 2012. doi: 10.4415/ANN\_12\_03\_06.



40. GOMES, J. H. F. *Método dos polinômios canônicos de misturas para otimização multi-objetivo*. Itajubá, Minas Gerais, Brasil: Tese de Doutorado - Programa de Pós Graduação em Engenharia de Produção - Universidade Federal de Itajubá, 2013.
41. GOPALAKRISHNAN, H.; KOSANOVIC, D. *Computers & Operations Research Operational planning of combined heat and power plants through genetic algorithms for mixed 0 – 1 nonlinear programming*. Computers and Operation Research, v. 56, p. 51–67, 2015. Elsevier.
42. Grzadzinski R, Huerta M, Lord C (2013). DSM-5 and autism spectrum disorders (ASDs): an opportunity for identifying ASD subtypes. *Mol Autism* 4: 12.
43. HAFER, J. F. LENHOFF, M. W. SONG, J. JORDAN J. M. HANNAN, M. T. HILLSTROM, H. J. *Reliability of plantar pressure platforms*. *Gait & posture*. 2013. doi: 10.1016/j.gaitpost.2013.01.028.
44. HASSAN, R. MOHAMMED, A. JANATI, I. YOUSSEF, G. MOHAMED, E. *Multilayer Perceptron: Architecture Optimization and Training*. International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 4, Nº1. 2016.
45. HAYKIN, S. *Redes Neurais: Princípios e prática*. Porto Alegre: Bookman, 2001.
46. HILTON, C. L. ZHANG, Y. WHILTE, M. R. KLOHR, C. L. & CONSTANTINO, J. *Motor impairment in sibling pairs concordant and discordant for autism spectrum disorders*. *Autism*. 16(4), 430-441. 2011.
47. HOLLAND, J. H. *Adaptation in Natural and Artificial Systems*. 1975. University of Michigan Press.
48. HOSSEINI, S. A.; NIAEI, A.; SALARI, D.; NABAVI, S. R. *Modeling and optimization of combustion process of 2-propanol over perovskite-type LaMnyCo1-yO3 nanocatalysts by an unreplicated experimental design with mixture-process variables and genetic algorithm methodology*. Journal of the Taiwan Institute of Chemical Engineers, v. 45, n. 1, p. 85–91, 2014. Taiwan Institute of Chemical Engineers.
49. ITURRIAGA, F.J.L.; SANZ, I. P. *Bankruptcy visualization and prediction using neural networks: a study of US commercial banks*, *Exp. Syst. Appl.* 42, p. 2857–2869, 2015.
50. JESTE, S. and S, GESCHWIND, D. H. *Disentangling the heterogeneity of autism spectrum disorder through genetic findings*. *Nat Rev Neurol*. 2014. 10: 74-81.
51. KABIR, M. M.; SAINT, D. A.; NALIVAICO, E.; ABBOTT, D.; VOSS, A.; BAUMERT, M. *Quantification of Cardiorespiratory Interactions Based on Joint Symbolic Dynamics*. *Annals of Biomedical Engineering*. 2011.
52. KINDREGAN, D. GALLAGHER, L. GORMLEY, J. *Gait Deviations in Children with Autism Spectrum Disorders: A Review*. *Autism Research and Treatment*. 2015.
53. KRISTJANPOLLER, W.; MINUTOLO, M. C. *Gold price volatility: a forecasting approach using the Artificial Neural Network–GARCH model*, *Exp. Syst. Appl.* 2015.
54. LABOISSIERE, L. A.; FERNANDES, R. A.; LAGE, G. G. *Maximum and minimum stock price forecasting of Brazilian power distribution companies based on artificial neural networks*, *Appl. Soft Comput.* 2015.
55. LARRAIN, J. M.; TURNER, C. *The treasury bill rate, the great recession, and neural networks estimates of real business sales*, *Proc. Comput. Sci.* 36, p. 227–233, 2014.

56. LI, J.; YANG, X.; REN, C.; CHEN, G.; WANG, Y. *Multiobjective optimization of cutting parameters in milling process using nondominated sorting genetic algorithm*. The International Journal of Advanced Manufacturing Technology, v. 76, p. 941–953, 2014.
57. LIM, B. DAVID, O. CHOI, B. KIM, M. *Comparative gait analysis between children with autism and age-matched controls: analysis with temporal-spatial and foot pressure variables*. The Journal of Physical Therapy Science. 2016.
58. LIM, T. Y.; AL-BETAR, M. A.; KHADER, A. T. *Adaptive pair bonds in genetic algorithm : An application to real-parameter optimization*. Applied Mathematics and Computation, v. 252, p. 503–519, 2015.
59. LIM, B. DAVID. S. CHOI. B. KIM. M. *Comparative gait analysis between children with autism and age-matched controls: analysis with temporal-spatial and foot pressure variables*. The Journal of Physical Therapy Science. V28. 286–292, 2016.
60. MALAQUIAS, N. G. L. *Uso dos Algoritmos Genéticos para a Otimização de Rotas de Distribuição*. Dissertação de mestrado em Ciências. Universidade Federal de Uberlândia. Uberlândia – Minas Gerais. 2006.
61. MARTINEZ-NOVA, A., J. C. CUEVAS-GARCÍA, J. PASCUAL-HUERTA. (2007) - "BioFoot in-shoe system: Normal values and assessment of the reliability and repeatability". The Foot.
62. MARTINIANO, A.; FERREIRA, R. P.; FERREIRA, A.; FERREIRA, A.; SASSI, R.J. *utilizando uma rede neural artificial para aproximação da função de evolução do sistema de lorentz*. Revista Produção e Desenvolvimento, v.2, n.1, p.26-38, jan./abr., 2016.
63. MCCARY, L. M. & J. E. Roberts. *Early identification of autism in fragile X syndrome: a review*. Journal of Intellectual Disability Research. volume 57 part 9 pp 803–814 september 2013. doi: 10.1111/j.1365-2788.2012.01609.x
64. MEIRELES, T. M. *Análise Biomecânica do Padrão de Marcha nas Perturbações do Espectro Autista*. Dissertação de Mestrado em Ciências do Desporto – Área de Especialização em Atividade Física Adaptada. Universidade do Porto. Porto-Portugal. 2014.
65. MEMARI, A. H. GHANOUNI, P. GHARIBZADEH, S. EGHLIDI, J. ZIAEE, V. & MOSHAYEDI, P. *Postural sway patterns in children with autism spectrum disorder compared with typically developing children*. Research in Autism Spectrum Disorders. 7, 325-332. 2013.
66. MENEZES, P. L. AZEVEDO, C. A. V. EYNG, E. NETO, J. D. LIMA, V. L. A. *Artificial neural network model for simulation of water distribution in sprinkle irrigation*. Rev. bras. eng. agríc. ambient. vol.19 no.9 Campina Grande Sept. 2015.
67. MERCK. 2011. Retirado: [www.manualmerck.net/images/thumbnail/p\\_347.gif](http://www.manualmerck.net/images/thumbnail/p_347.gif).
68. MORAIS, E. C. *Reconhecimento de padrões e redes neurais artificiais em predição de estruturas secundárias de proteínas*. Tese de doutorado em Engenharia de Sistemas e Computação. COPPE - Universidade Federal do Rio de Janeiro. 2010.
69. NAYATE, A. TONGE, B. J. BRADSHAW, J. L. MCGINLEY, J. L. IANSEK, R. & RINEHART, N. J. *Differentiation of High-Functioning Autism and Asperger's Disorder Based on Neuromotor Behaviour*. J Autism Dev Disord. 42, 707-717. 2012.

70. NAYATE, A., Bradshaw, J., Rinehart, N. *Autism and asperger's disorder: Are they movement disorders involving the cerebellum and/or basal ganglia?* Brain Research Bulletin, 67, 327-334. 2005.
71. NOBILE. M.PEREGO. P. PICCININI. L. MANI. E. ROSSI. A. BELLINA. M. MOLteni. M. *Further evidence of complex motor dysfunction in drug naïve children with autism using automatic motion analysis of gait.* Autism. SAGE Publications and The National Autistic Society Vol 15(3). 2011.
72. PAPPAS, E. FERNANDO, M. E. CROWTHER, R.G. LAZZARINI, P. A. CUNNINGHAM, M. SANGLA, K. S. *Plantar Pressure in Diabetic Peripheral Neuropathy Patients with Active Foot Ulceration, Previous Ulceration and No History of Ulceration: A Meta-Analysis of Observational Studies.* PLoS ONE. 2014. Doi: 10.1371/journal.pone.0099050
73. PAULA, T. I. *Avaliação da influência de parâmetros do Algoritmo Genético na otimização de um problema multiobjetivo utilizando-se Arranjo de Misturas.* Dissertação de mestrado em Ciências em Engenharia de Produção. Universidade Federal de Itajubá. 2015.
74. PEDRO, S. FRANCO, A. BORELLA, C. P. ROCHA, E. S. CARPESA, F. P. *Análise da variabilidade e repetibilidade da pressão plantar durante a marcha de idosos.* Rev. Bras. Reumatol. vol.55 no.5 São Paulo Sept./Oct. 2015
75. PEREIRA, D. M. *Análise dos efeitos de um plano educacional individualizado no desenvolvimento acadêmico e funcional de um aluno com transtorno do espectro do autismo.* Dissertação de mestrado, Universidade Federal do Rio Grande do Norte, Natal, RN, Brasil. 2014.
76. PERERA, A. CRUZ, A. BONATTO, C. RECH, F. SIMONAGGIO, L. GIACOMET, D e SCHUSTER, R. *Análise do padrão de marcha do espectro autista.* II Congresso de Pesquisa e Extensão da FSG. ISSN 2318-8014. 2014.
77. PINEDA, J. O. C. *Entropia e Teoria da Informação: Os fundamentos Científicos da Era Digital.* 1. ed. São Paulo, 2010.
78. PINTO, V. M. N. *Caracterização da pressão plantar em amputados transtibiais e transfemorais.* Dissertação de mestrado em engenharia de reabilitação e acessibilidade humanas. Universidade de trás-os-montes e alto douro. Vila Real - Portugal. 2014.
79. PIRES, T. S. *Método de superfície de resposta aplicado à otimização termoeconômica de sistemas de cogeração modelados em um simulador de processos.* Dissertação de Mestrado em Engenharia Mecânica. COPPE – Universidade Federal do Rio de Janeiro. Rio de Janeiro. 2010.
80. QIU, X. TIAN, D. H. HAN, C. L. CHEN, W. WANG, Z. J. UM, Z. Y. *Risk Factors Correlated with Plantar Pressure in Chinese Patients with Type 2 Diabetes.* Diabetes Technol Ther. 2013.
81. RATH, S.; JAGADEV, A. K.; NAYAK, M. R. *Performance analysis of stock market using artificial neural network,* Int. J. Appl. Eng. Res. 10, 2015.
82. RAZNAHAN A, WALLACE GL, ANTEZANA L. *Early brain overgrowth in autism and the perils of population norms.* Biol Psychiatry. 2013;74(8):563–575

83. REGO, S. W. S. E. *Autismo: Fisiopatologia e biomarcadores*. Dissertação de mestrado em medicina. Universidade da Beira Interior. Cavilhã - Portugal. 2012.
84. RETICO, A. TOSETTI, M. MURATORI, F. CALDERONI S. *Neuroimaging-based methods for autism identification: a possible translational application?*. *Functional Neurology*. 2014. 29(4): 231-239.
85. RIBEIRO, I. P. FREITAS, M. TELES, N. *As Perturbações do Espectro do Autismo – Avanços da Biologia Molecular*. *Nascer e Crescer*, vol. 22, 2013.
86. ROBERT, M. LAURENT, M. AXEL, U. KENNETH, M. DANIEL, T. *Plantar pressure measurements and running-related injury: A systematic review of methods and possible associations*. *Gait & Posture*. Volume 47, June 2016.
87. ROCHA, C. S. C. *Técnicas de Reconhecimento de Padrões para a Classificação de Sinais Fisiológicos*. Tese de doutorado em Engenharia Elétrica. Universidade Federal de Minas Gerais. 2013.
88. RUSSEL, S.; NORVIG, P. *Inteligência Artificial*. Campus. 3. ed. São Paulo, 2013.
89. SAITO, M., K. NAKAJIMA, C. TAKANO (2011) - "An in-shoe device to measure plantar pressure during daily human activity". *Medical Engineering & Physics*.
90. SANTOS, C. P. E. F. *Dinâmica simbólica na análise de apneuse e de tarefa sensório-motora em ressonância magnética funcional*. Dissertação de Mestrado em Engenharia Biomédica. Universidade Federal do Rio de Janeiro. 2008.
91. SANTOS, J. D. dos, SILVA, Z. C. *Métodos Numéricos*, 3ª ed., Recife. Editora Universitária UFPE, 2010.
92. SANTOS, M. C.; MACHADO, A. R.; BARROZO, M. A. S.; JACKSON, M. J.; EZUGWU, E. O. *Multi-objective optimization of cutting conditions when turning aluminum alloys using genetic algorithm*. *The International Journal of Advanced Manufacturing Technology*, v. 76, p. 1123–1138, 2014.
93. SANTOS, M. *Uso da Entropia de Shannon para classificação de zonas artificiais ou naturais em imagens*. Dissertação de Mestrado de Engenharia Informática. Universidade da Beira Interior. Covilhã - Portugal. 2013.
94. SEITZ, P. (2008) - "Novel.de." Retrieved 02/04/2010, 2011, from <http://novel.de/novelcontent/software/emed/emedd>.
95. SOARES, F. C.; ROBAINA, A. D.; PEITER, M. X.; RUSSI, J. L.; VIVIAN, G. A. *Redes neurais artificiais na estimativa da retenção de água do solo*. *Ciência Rural*, v.44, p.293-300, 2014.
96. SOARES, I. *Tratamento de dados de intensidade de sinal celular utilizando interpolação*. 5th International Information and Telecommunication Technologies Symposium. Cuiabá-MT 2006.
97. SOUSA, A. P. F. *Um algoritmo Genético para o planeamento de rotas com considerações ambientais*. Dissertação de Mestrado. Modelação, análise de dados e sistemas de apoio à decisão otimização. Universidade do Porto. 2015.
98. SOUSA, Daniela Sofia S., and João Manuel R. S. Tavares. 2006. *Estudo de Soluções Comerciais para Análise Clínica da Pressão Plantar Porto*: Faculdade de Engenharia da Universidade do Porto.

99. SOUZA, L. L. (2012). *Caracterização e detecção automática de eventos epileptiformes em sinais de eletroencefalografia por dinâmica simbólica*. Dissertação de Mestrado em Ciências Mecânicas, Publicação MTARH.DM, Departamento de Engenharia Mecânica, Universidade de Brasília, Brasília, DF, 81 p.
100. SOUZA, R. P. *Otimização de treliças com restrições de falha combinando técnicas de programação de algoritmos contínuos e discretos*. Dissertação de mestrado em Engenharia. Universidade Federal do Rio Grande do Sul. Porto Alegre – RS. 2009.
101. SUBASIA, A. & ERCELEBI, E. *Classification off eeg signals using neural network and logistic regression*. Computer Methods and Programs in Biomedicine, vol. 78, no. , May 2005.
102. TÁBUAS, C. *Análise da pressão plantar para fins de diagnóstico*. Dissertação de Mestrado em Engenharia Biomédica. Faculdade de Engenharia da Universidade do Porto. Portugal. 2012.
103. THIESEN, T. SUMIYA, A. *Equilíbrio e arco plantar no balé clássico*. ConScientiae Saúde - 10: 138-142. 2011.
104. TSCHINKEL, P. F. S. *Avaliação das concentrações plasmáticas dos elementos traço cobre, zinco e selênio em crianças com transtorno do espectro autista*. Dissertação de mestrado em Saúde e Desenvolvimento na Região Centro-Oeste. Universidade Federal do Mato Grosso do Sul. Campo Grande – MS. 2014.
105. VALENCIA, F. J; VALLVERDU, M.; RIVERO, I.; VOSS A.; BAYES, A. L.; PORTA, A.; CAMINAL, P. *Symbolic dynamics to discriminate healthy and ischemic dilated cardiomyopathy populations: an application to the variability of heart period and QT interval*. Philosophical Transaction Royal Society. A373:20140092. 2015.
106. VICTORIA L. C CALHOUN. M. *Symmetry of gait in children with autism*. Autism Research and Treatment. 2012. Article ID 576478.
107. WEISS, M. J. MORAN, M. F. PARKER, M. E. FOLEY. J. T. *Gait analysis of adolescents and young adults diagnosed with autism and severe disorders of verbal communication*. Frontiers in integrative neuroscience. 2013.
108. XIMENES, Mileny; Peron, Guilherme.2013. Implementação e avaliação de uma palmilha de látex com sistema de sensores para investigação da distribuição de pressão plantar em indivíduos autistas. Dissertação de graduação em engenharia eletrônica, Universidade de Brasília, Campus Gama, DF, 77p.
109. ZWAIGENBAUM, L. BRYSONB, S. GAROND, N. *Early identification of autism spectrum disorders*. Behavioural Brain Research. 2013.
110. ZWAIGENBAUM, L. MARGARET, BAUMAN, L. STONE, C. ROBIN, L. *Early Identification of Autism Spectrum Disorder: Recommendations for Practice and Research*. Pediatrics. 2015. 136:S10–S40.