

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
ESCOLA DE ENGENHARIA
DEPARTAMENTO DE ENGENHARIA ELÉTRICA
ENGENHARIA DE COMPUTAÇÃO

ARTHUR CRIPPA BÚRIGO

**Classificação de Movimentos da Mão
Utilizando Eletromiografia de Superfície,
Regressão Logística, Redes Neurais,
Máquina de Vetores de Suporte e Base de
Dados NinaPro**

Monografia apresentada como requisito parcial
para a obtenção do grau de Bacharel em
Engenharia de Computação.

Prof. Dr. Alexandre Balbinot
Orientador

Porto Alegre
2014

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Vice-Reitor: Prof. Rui Vicente Oppermann

Pró-Reitor de Graduação: Prof. Sérgio Roberto Kieling

Diretor do Instituto de Informática: Prof. Luís da Cunha Lamb

Diretor da Escola de Engenharia: Prof. Luiz Carlos Pinto da Silva Filho

Coordenador do Curso de Engenharia de Computação: Prof. Marcelo Götz

Coordenador do Curso de Engenharia Elétrica: Prof. Ramon Carlos Poisl

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

SUMÁRIO

LISTA DE ABREVIATURAS E SIGLAS	6
LISTA DE FIGURAS	7
LISTA DE TABELAS	9
ABSTRACT	16
RESUMO	17
1 INTRODUÇÃO	18
1.1 Justificativa	18
1.2 Objetivos	19
2 SINAIS MIOELÉTRICOS	20
2.1 Contração e Relaxamento Muscular	20
2.1.1 Neurônio	21
2.1.2 Potencial de Ação (Impulso Nervoso)	22
2.1.3 Junção Neuromuscular	23
2.1.4 Músculo Esquelético	24
2.2 Eletromiografia	26
2.3 Musculatura do Antebraço	28
3 BASE DE DADOS NINAPRO	32
3.1 Protocolo de Aquisição de Dados	32
3.2 Conjunto de Dados	35
3.3 Conjunto de Poses Seleccionadas	36
4 PRÉ-PROCESSAMENTO	37
4.1 Ajuste de <i>Offset</i>	37
4.2 Filtragem	37
4.3 Normalização	39
4.4 Segmentação	40
4.5 Extração de Características	41
4.5.1 RMS	42
4.5.2 Média Aritmética	42
4.5.3 Variância	42
4.5.4 Desvio Padrão	42
4.5.5 Obliquidade	42

4.5.6	Curtose	43
4.5.7	Centroide Espectral	43
4.6	<i>Feature Scaling</i>	43
5	APRENDIZADO DE MÁQUINA	44
5.1	Regressão Logística	44
5.1.1	Modelo	44
5.1.2	Regularização	46
5.1.3	Predição	47
5.1.4	Classificação “Um Contra Todos”	47
5.2	Redes Neurais	47
5.2.1	<i>Feedforward</i>	48
5.2.2	Treinamento	49
5.2.3	Retropropagação de Erros	50
5.2.4	Regularização	52
5.2.5	Predição	53
5.2.6	Valores Iniciais para os Pesos	53
5.2.7	Quantidade de Camadas e de Unidades em cada Camada	53
5.3	Máquina de Vetores de Suporte	54
5.3.1	Modelo	54
5.3.2	Função <i>Kernel</i>	59
5.3.3	Predição	60
5.3.4	Classificação “Um Contra Todos”	60
5.4	Validação Cruzada	61
5.5	Métricas de Qualidade de Classificação	61
5.5.1	Matriz de Confusão	61
5.5.2	Taxa de Acerto	62
5.5.3	Valor Preditivo Positivo	62
5.5.4	Taxa de Verdadeiros Positivos	62
5.5.5	Taxa de Verdadeiros Negativos	63
5.5.6	Valor Preditivo Negativo	63
5.5.7	<i>F1 Score</i>	63
5.5.8	<i>Informedness</i>	64
5.5.9	<i>Markedness</i>	64
6	RESUMO DE TRABALHOS ANTERIORES	65
7	METODOLOGIA EXPERIMENTAL	67
7.1	Plataforma de Software	67
7.2	Funcionamento do Sistema	67
7.2.1	Leitura dos Dados de Entrada de um Sujeito	68
7.2.2	Teste de Algoritmo de Aprendizado	70
7.2.3	Conjuntos de Treinamento e de Teste	71
7.3	Determinação de Parâmetros dos Modelos	73
8	RESULTADOS E DISCUSSÕES	74
8.1	Desempenho Variando o Conjunto de Eletrodos	74
8.1.1	Utilizando 8 Eletrodos em Volta do Antebraço e 4 Eletrodos em Músculos Selecionados	74

8.1.2	Utilizando 4 Eletrodos em Volta do Antebraço e 4 Eletrodos em Músculos Seleccionados	76
8.1.3	Utilizando 4 Eletrodos em Músculos Seleccionados	77
8.1.4	Utilizando 8 Eletrodos em Volta do Antebraço	79
8.1.5	Utilizando 4 Eletrodos em Volta do Antebraço	80
8.1.6	Utilizando 2 Eletrodos em Volta do Antebraço	82
8.1.7	Análise de Desempenho Variando o Conjunto de Eletrodos	83
8.2	Desempenho Variando o Conjunto de Características	86
8.2.1	Utilizando Todas as Características	87
8.2.2	Ignorando o Valor RMS	88
8.2.3	Ignorando a Média Aritmética	89
8.2.4	Ignorando a Variância	91
8.2.5	Ignorando o Desvio Padrão	92
8.2.6	Ignorando a Obliquidade	94
8.2.7	Ignorando a Curtose	95
8.2.8	Ignorando o Centróide Espectral	97
8.2.9	Ignorando o Domínio do Tempo	98
8.2.10	Ignorando o Domínio da Frequência	100
8.2.11	Somente o Valor RMS no Domínio do Tempo	101
8.2.12	Somente a Média Aritmética no Domínio do Tempo	103
8.2.13	Somente a Variância no Domínio do Tempo	105
8.2.14	Somente o Desvio Padrão no Domínio do Tempo	106
8.2.15	Somente a Obliquidade no Domínio do Tempo	108
8.2.16	Somente a Curtose no Domínio do Tempo	110
8.2.17	Somente o Valor RMS no Domínio da Frequência	111
8.2.18	Somente a Média Aritmética no Domínio da Frequência	113
8.2.19	Somente a Variância no Domínio da Frequência	115
8.2.20	Somente o Desvio Padrão no Domínio da Frequência	116
8.2.21	Somente a Obliquidade no Domínio da Frequência	118
8.2.22	Somente a Curtose no Domínio da Frequência	119
8.2.23	Somente o Centróide Espectral no Domínio da Frequência	121
8.2.24	Análise de Desempenho Variando o Conjunto de Características	122
8.3	Desempenho Variando o Tamanho do Conjunto de Treinamento	127
8.3.1	Utilizando 90% dos Exemplos para o Treinamento	127
8.3.2	Utilizando 80% dos Exemplos para o Treinamento	128
8.3.3	Utilizando 50% dos Exemplos para o Treinamento	130
8.3.4	Utilizando 40% dos Exemplos para o Treinamento	131
8.3.5	Utilizando 20% dos Exemplos para o Treinamento	133
8.3.6	Análise de Desempenho Variando o Tamanho do Conjunto de Treinamento	134
9	CONCLUSÕES	137
	REFERÊNCIAS	138
	ANEXO A FUNÇÃO DE MINIMIZAÇÃO POR CARL EDWARD RASMUSSEN	141

LISTA DE ABREVIATURAS E SIGLAS

EMG	<i>Electromyography</i> (Eletromiografia)
FFLS	<i>Finger-Force Linear Sensor</i>
KRLS	<i>Kernel Regularized Least Squares</i>
LDA	<i>Linear Discriminant Analysis</i> (Análise Discriminante Linear)
MU	<i>Motor Unit</i> (Unidade Motora)
MUAP	<i>Motor Unit Action Potential</i> (Potencial de Ação de Unidade Motora)
MVA	Média do Valor Absoluto
NinaPro	<i>Non-Invasive Adaptive Prosthetics</i>
RBF	<i>Radial Basis Function</i> (Função de Base Radial)
RMS	<i>Root Mean Square</i> (Valor Quadrático Médio)
sEMG	<i>Superficial Electromyography</i> (Eletromiografia Superficial)
SMO	<i>Sequential Minimal Optimization</i> (Otimização Mínima Sequencial)
SVM	<i>Support Vector Machine</i> (Máquina de Vetores de Suporte)

LISTA DE FIGURAS

2.1	A estrutura de um neurônio (adaptado de TORTORA; DERRICKSON, 2013).	22
2.2	Potencial de Ação ou Impulso (adaptado de TORTORA; DERRICKSON, 2013).	23
2.3	Estrutura da junção neuromuscular, a sinapse entre um neurônio motor somático e uma fibra muscular esquelética (adaptado de TORTORA; DERRICKSON, 2013).	24
2.4	Organização do músculo esquelético, do nível macroscópico ao molecular. As letras F, G, H e I são cortes transversais nos níveis indicados (adaptado de GUYTON; HALL, 2010).	25
2.5	Relação entre músculos esqueléticos e ossos. Os ossos servem como alavancas, e as articulações servem como fulcros para as alavancas (adaptado de TORTORA; DERRICKSON, 2013).	26
2.6	Representação esquemática da geração do sinal mioelétrico (adaptado de RICCIOTTI, 2006).	27
2.7	Músculos do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos. Da esquerda para a direita, visões anterior superficial, anterior intermediária e anterior profunda (adaptado de TORTORA; DERRICKSON, 2013).	29
2.8	Músculos do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos. Da esquerda para a direita, visões posterior superficial e posterior profunda (adaptado de TORTORA; DERRICKSON, 2013).	31
3.1	Posicionamento dos doze eletrodos superficiais no braço. O eletrodo do músculo flexor superficial dos dedos está oculto pelo braço e, portanto, não é visível na imagem (adaptado de GIJSBERTS <i>et al.</i> , 2014).	33
3.2	Os 49 movimentos (ou posturas) mais posição de repouso do segundo banco de dados NinaPro (adaptado de GIJSBERTS <i>et al.</i> , 2014).	34
3.3	Os 6 movimentos (ou posturas) selecionados do segundo banco de dados NinaPro (adaptado de GIJSBERTS <i>et al.</i> , 2014).	36
4.1	Exemplo de ajuste de <i>offset</i> do sinal representado pela Equação (4.2).	38
4.2	Resposta em frequência do filtro digital passa-banda com frequências de corte 20 Hz e 500 Hz de quarta ordem e com resposta Butterworth.	38
4.3	Espectro de frequência do sinal representado pela Equação (4.4) antes e após a filtragem.	39
4.4	Exemplo de normalização do sinal representado pela Equação (4.6).	40

4.5	Sinal representado pela Equação (4.11) antes e após a segmentação.	41
5.1	Função sigmoide.	45
5.2	Arquitetura da rede neural utilizada neste projeto.	48
5.3	Função sigmoide (a) e seu gradiente (b).	52
5.4	A fronteira de decisão e o conceito de margem. Ao se maximizar a margem, obtém-se uma fronteira de decisão cuja localização é determinada por um subconjunto de pontos, conhecidos como vetores de suporte (indicados pelos círculos) (BISHOP, 2006).	55
5.5	Variáveis de folga $\xi_n \geq 0$. Pontos com círculos são vetores de suporte (BISHOP, 2006).	58
5.6	Núcleo gaussiano (ou função de base radial) com $\sigma = 1$	60
7.1	Diagrama de fluxo do comportamento geral do sistema.	68
7.2	Diagrama de fluxo da leitura dos dados de um sujeito.	69
7.3	Diagrama de fluxo do pré-processamento de um sinal.	70
7.4	Diagrama de fluxo do teste de um algoritmo de aprendizado.	71
7.5	Divisão do conjunto de exemplos entre os conjuntos de treinamento (verde) e de teste (vermelho).	72
7.6	Seis repetições em cada movimento.	72
7.7	Divisão dos exemplos de um mesmo movimento entre os conjuntos de treinamento e de teste.	73
8.1	Taxa de acerto variando o conjunto de eletrodos.	84
8.2	<i>Informedness</i> variando o conjunto de eletrodos.	85
8.3	Taxa de acerto variando o conjunto de características (ignorando características).	123
8.4	Taxa de acerto variando o conjunto de características (utilizando somente uma das características).	125
8.5	<i>Informedness</i> variando o conjunto de características (utilizando somente uma das características).	126
8.6	Taxa de acerto variando o tamanho do conjunto de treinamento.	135
8.7	F1 score variando o tamanho do conjunto de treinamento.	136

LISTA DE TABELAS

2.1	Músculos do compartimento anterior do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos e suas funções específicas (adaptado de TORTORA; DERRICKSON, 2013).	28
2.2	Músculos do compartimento posterior do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos e suas funções específicas (adaptado de TORTORA; DERRICKSON, 2013).	30
5.1	Matriz de confusão para o caso binário.	62
6.1	Relação entre o método de classificação utilizado e a taxa de erros do classificador (GLETTE <i>et al.</i> , 2008).	65
6.2	Relação entre a quantidade de movimentos considerados e a taxa de acerto do classificador (BOSCHMANN <i>et al.</i> , 2009).	66
6.3	Relação entre a quantidade de movimentos considerados e a taxa de erros do classificador (ATZORI <i>et al.</i> , 2012).	66
6.4	Relação entre o núcleo (<i>kernel</i>) de classificação utilizado, o conjunto de características utilizado e a taxa de acerto do classificador (GIJSBERTS <i>et al.</i> , 2014).	66
8.1	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	75
8.2	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	75
8.3	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	76
8.4	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	76
8.5	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	77
8.6	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.	77

8.7	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.	78
8.8	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.	78
8.9	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.	79
8.10	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.	79
8.11	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.	80
8.12	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.	80
8.13	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.	81
8.14	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.	81
8.15	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.	82
8.16	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando utilizando 2 eletrodos em volta do antebraço. . .	82
8.17	Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.	82
8.18	Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.	83
8.19	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.	83
8.20	Desempenho do modelo de regressão logística contra o conjunto de teste quando considerando todas as características.	87
8.21	Desempenho do modelo de rede neural contra o conjunto de teste quando considerando todas as características.	87
8.22	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando considerando todas as características.	88
8.23	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o valor RMS do sinal.	88
8.24	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o valor RMS do sinal.	89
8.25	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o valor RMS do sinal.	89
8.26	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a média aritmética do sinal.	90
8.27	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a média aritmética do sinal.	90
8.28	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a média aritmética do sinal. . . .	91
8.29	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a variância do sinal.	91

8.30	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a variância do sinal.	92
8.31	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a variância do sinal.	92
8.32	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o desvio padrão do sinal.	93
8.33	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o desvio padrão do sinal.	93
8.34	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o desvio padrão do sinal.	94
8.35	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a obliquidade do sinal.	94
8.36	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a obliquidade do sinal.	95
8.37	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a obliquidade do sinal.	95
8.38	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a curtose do sinal.	96
8.39	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a curtose do sinal.	96
8.40	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a curtose do sinal.	97
8.41	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o centroide espectral do sinal.	97
8.42	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o centroide espectral do sinal.	98
8.43	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o centroide espectral do sinal.	98
8.44	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.	99
8.45	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.	99
8.46	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.	100
8.47	Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.	100
8.48	Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.	101
8.49	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.	101
8.50	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o valor RMS do sinal no domínio do tempo.	102

8.51	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.	102
8.52	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.	102
8.53	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.	103
8.54	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a média aritmética do sinal no domínio do tempo.	103
8.55	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.	104
8.56	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.	104
8.57	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.	105
8.58	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a variância do sinal no domínio do tempo.	105
8.59	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.	105
8.60	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.	106
8.61	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.	106
8.62	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o desvio padrão do sinal no domínio do tempo.	107
8.63	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.	107
8.64	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.	107
8.65	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.	108
8.66	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a obliquidade do sinal no domínio do tempo.	108

8.67	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.	109
8.68	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.	109
8.69	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.	110
8.70	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a curtose do sinal no domínio do tempo.	110
8.71	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.	110
8.72	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.	111
8.73	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.	111
8.74	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o valor RMS do sinal no domínio da frequência.	112
8.75	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.	112
8.76	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.	112
8.77	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.	113
8.78	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a média aritmética do sinal no domínio da frequência.	113
8.79	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.	114
8.80	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.	114
8.81	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.	115
8.82	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a variância do sinal no domínio da frequência.	115

8.83	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.	115
8.84	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.	116
8.85	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.	116
8.86	Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o desvio padrão do sinal no domínio da frequência.	117
8.87	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.	117
8.88	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.	117
8.89	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.	118
8.90	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.	118
8.91	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.	119
8.92	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.	119
8.93	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.	120
8.94	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.	120
8.95	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.	121
8.96	Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o centroide espectral do sinal no domínio da frequência.	121
8.97	Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o centroide espectral do sinal no domínio da frequência.	122
8.98	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o centroide espectral do sinal no domínio da frequência.	122

8.99	Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.	127
8.100	Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.	128
8.101	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.	128
8.102	Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.	129
8.103	Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.	129
8.104	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.	130
8.105	Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.	130
8.106	Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.	131
8.107	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.	131
8.108	Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.	132
8.109	Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.	132
8.110	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.	133
8.111	Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.	133
8.112	Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.	134
8.113	Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.	134

Classification of Hand Movements Using Superficial Electromyography, Logistic Regression, Neural Network, Support Vector Machine and NinaPro Database

ABSTRACT

People with disabilities are partially limited due to their physical condition. One way to minimize their situation is the use of prosthesis to reduce or, ideally, completely eliminate the disability. This project uses the second NinaPro database to develop a system, based on superficial electromyography, capable of classifying movements of a person's hand (abduction of the fingers, fingers flexed together in fist, pointing index, wrist flexion, wrist extension and grasp). In particular, it compares the accuracy of three classification methods (logistic regression, neural networks and support vector machine) for variable sets of electrodes, for variable sets of characteristics and for training sets of different sizes. This paper demonstrates that the accuracies of the models decrease with the reduction of the set of electrodes, but shows that the positioning of these electrodes may be an important factor for the quality of classification. Moreover, it highlights high accuracy rates when considering only some features extracted from the frequency domain (as skewness and spectral centroid), confirms that the accuracies of the models decrease with the reduction of the training set size and shows that logistic regression and neural network models present similar but superior accuracies (96,3 % on the standard test case), when compared to that of a support vector machine model (92,2 % on the standard test case).

Keywords: Superficial Electromyography, Second NinaPro Database, Hand Movement Classification, Logistic Regression, Neural Network, Support Vector Machine.

RESUMO

Portadores de deficiência têm parte de seu poder de ação limitado devido a sua condição física. Uma maneira de amenizar a situação dessas pessoas é a utilização de próteses para a eliminação parcial ou, idealmente, completa da deficiência. Este trabalho utiliza a segunda base de dados NinaPro para o desenvolvimento de um sistema, baseado em eletromiografia de superfície, para a classificação de movimentos da mão de uma pessoa (abdução dos dedos, dedos flexionados em punho, extração do dedo indicador, flexão da mão, extensão da mão e pegada). Em especial, ele compara o desempenho de três algoritmos de classificação (regressão logística, redes neurais e máquina de vetores de suporte) para conjuntos variáveis de eletrodos, para conjuntos variáveis de características e para tamanhos de conjunto de treinamento distintos. O trabalho mostra que as taxas de acerto dos modelos de aprendizado decrescem com a diminuição do tamanho do conjunto de eletrodos utilizado, porém evidencia que o posicionamento destes eletrodos também pode ser um fator importante para a qualidade da classificação dos sinais. Além disso, ele destaca altas taxas de acerto ao se considerar somente algumas características extraídas do domínio da frequência (como a obliquidade e o centroide espectral), confirma que a taxa de acerto média dos modelos diminui com a redução do tamanho do conjunto de treinamento e revela que os modelos de regressão logística e rede neural apresentam taxas de acerto muito próximas (96,3 % no caso de teste padrão) e superiores às taxas do modelo de máquina de vetores de suporte (92,2 % no caso de teste padrão).

Palavras-chave: Eletromiografia de Superfície, Segunda Base de Dados NinaPro, Classificação de Movimentos da Mão, Regressão Logística, Redes Neurais, Máquina de Vetores de Suporte.

1 INTRODUÇÃO

Pessoas portadoras de deficiência são, de um modo geral, impossibilitadas de exercer ou participar, independentemente, de muitas atividades comuns do dia-a-dia de qualquer pessoa. Devido a sua situação, elas acabam sendo excluídas e, até mesmo, excluindo a si mesmas da sociedade em que vivem. Por esta e outras razões, estuda-se, constantemente, formas de incluir, socialmente, estas pessoas (ATKINS; HEARD; DONOVAN, 1996; ATZORI *et al.*, 2012).

Uma das soluções possíveis é a eliminação parcial ou, idealmente, completa de uma deficiência física pela utilização de próteses. Através da captação e interpretação de sinais elétricos transmitidos pelo corpo humano, pretende-se desenvolver dispositivos capazes de substituir ou complementar funcionalidades de partes parcial ou completamente disfuncionais do corpo (ATZORI *et al.*, 2012; FAVIEIRO, 2012; NILSON, 2014).

Em função disso, o foco deste trabalho é o estudo da interpretação e classificação dos sinais mioelétricos produzidos pelos músculos do braço com o intuito de movimentar as mãos de uma pessoa.

1.1 Justificativa

A solução de utilização de próteses para a eliminação completa de uma deficiência humana é uma idealização que está longe da real capacidade dos dispositivos atuais. Não é, então, sem razão que se estuda, constantemente, novos métodos para o controle e o desenvolvimento de tais equipamentos (ATKINS; HEARD; DONOVAN, 1996; BOSCHMANN *et al.*, 2009; ATZORI *et al.*, 2012).

O desenvolvimento de uma prótese capaz de substituir por completo a funcionalidade de uma parte do corpo humano – como, por exemplo, uma mão - é uma tarefa complexa, devido ao grande número de graus de liberdade e, conseqüentemente, ao grande número de movimentos possíveis que uma mão pode apresentar. Além disso, os movimentos corporais não são lineares no tempo (e muito menos discretos), o que traz uma outra dimensão ao problema enfrentado por estes dispositivos. Por estas e outras razões, a fidelidade de movimento das próteses existentes ainda não é suficiente (ATKINS; HEARD; DONOVAN, 1996; BOSCHMANN *et al.*, 2009; FAVIEIRO, 2009; FAVIEIRO; BALBINOT, 2011; FAVIEIRO; BALBINOT; BARRETO, 2011).

Sabe-se, devido a diversos estudos na área – como em BOSCHMANN *et al.* (2009), em MAZURKIEWICZ (2007) ou em ORTOLAN (2002) –, que pessoas amputadas ainda são capazes de gerar sinais musculares padronizados de intenção de movimento, que podem ser captados, interpretados e transformados no movimento real de uma prótese eletro-mecânica.

1.2 Objetivos

O objetivo deste trabalho é o desenvolvimento de um sistema, baseado em eletromiografia de superfície, que seja capaz de interpretar e classificar sinais musculares de intenção de movimento da mão (abdução dos dedos, dedos flexionados em punho, extração do dedo indicador, flexão da mão, extensão da mão e pegada). Em particular, o trabalho visa comparar o desempenho de três algoritmos de classificação (regressão logística, redes neurais e máquina de vetores de suporte) para conjuntos variáveis de eletrodos, para conjuntos variáveis de características e para tamanhos de conjunto de treinamento distintos.

2 SINAIS MIOELÉTRICOS

O sinal mioelétrico é um impulso elétrico que causa a contração das fibras musculares do corpo, provocando o movimento (ORTOLAN, 2002; FAVIEIRO, 2009, 2012; TORTORA; DERRICKSON, 2013).

Luigi Galvani foi um dos primeiros pesquisadores a estudar a bioeletricidade. Em suas pesquisas, ele descobriu que os músculos de sapos mortos se contraíam quando excitados por faíscas elétricas. Instigado pelo trabalho de Luigi, Carlo Matteucci, a partir de experimentos próprios, foi considerado o primeiro a medir uma corrente elétrica gerada por um tecido muscular. Anos mais tarde, Emil du Bois-Reymond foi capaz de descobrir o potencial de ação nervoso. Mais recentemente, as pesquisas sobre o assunto se intensificaram, levando ao surgimento de uma nova área, conhecida pelo nome de Eletromiografia (WHITTAKER, 1910; PEARCE, 2001; ORTOLAN, 2002; FAVIEIRO, 2009; BRESADOLA, 2011; FAVIEIRO, 2012).

2.1 Contração e Relaxamento Muscular

Segundo GUYTON; HALL (2010), o processo de contração muscular, envolve várias etapas, como propagação de potenciais de ação pelo nervo motor, transmissão sináptica, propagação dos potenciais pelas fibras musculares e, finalmente, contração muscular. Ou, mais precisamente, pode-se dizer que o início e a execução de uma contração muscular ocorrem nas seguintes etapas:

1. os potenciais de ação viajam pelo nervo motor até suas terminações nas fibras musculares;
2. em cada terminação, o nervo secreta pequena quantidade da substância neurotransmissora acetilcolina;
3. a acetilcolina age em uma área local da membrana da fibra muscular para abrir múltiplos canais “regulados pela acetilcolina” por meio de moléculas de proteína flutuando na membrana;
4. a abertura dos canais regulados pela acetilcolina permite a difusão de grande quantidade de íons sódio para o lado interno da membrana da fibra muscular. Isso desencadeia um potencial de ação na membrana;
5. o potencial de ação se propaga pela membrana da fibra muscular da mesma forma que potenciais de ação viajam pelas membranas das fibras nervosas;

6. o potencial de ação despolariza a membrana muscular, e grande parte da eletricidade do potencial de ação flui pelo centro da fibra muscular. Aí, ela causa a liberação de grandes quantidades de íons de cálcio pelo retículo sarcoplasmático;
7. os íons cálcio iniciam forças atrativas entre os filamentos de miosina e actina, fazendo que eles deslizem ao lado um do outro, o que é conhecido como “processo contrátil”;
8. após uma fração de segundo, os íons cálcio são bombeados de volta para o retículo sarcoplasmático por uma bomba de Ca^{+2} da membrana, onde permanecem armazenados até que um novo potencial de ação muscular se inicie; essa retirada dos íons cálcio das miofibrilas faz a contração muscular cessar.

Se dois estímulos são aplicados, um imediatamente após o outro, o músculo reponderá ao primeiro estímulo, mas não ao segundo. Quando uma fibra muscular recebe estímulo o suficiente para se contrair, ela perde, temporariamente, a excitabilidade e não pode responder por um tempo. O período de perda de excitabilidade, chamado de período refratário, é característico de todas as células musculares e nervosas. A duração do período refratário varia com o músculo envolvido. O músculo esquelético tem um período refratário pequeno de, mais ou menos, 5 ms; o músculo cardíaco tem um período refratário mais longo de, mais ou menos, 300 ms (TORTORA; DERRICKSON, 2013).

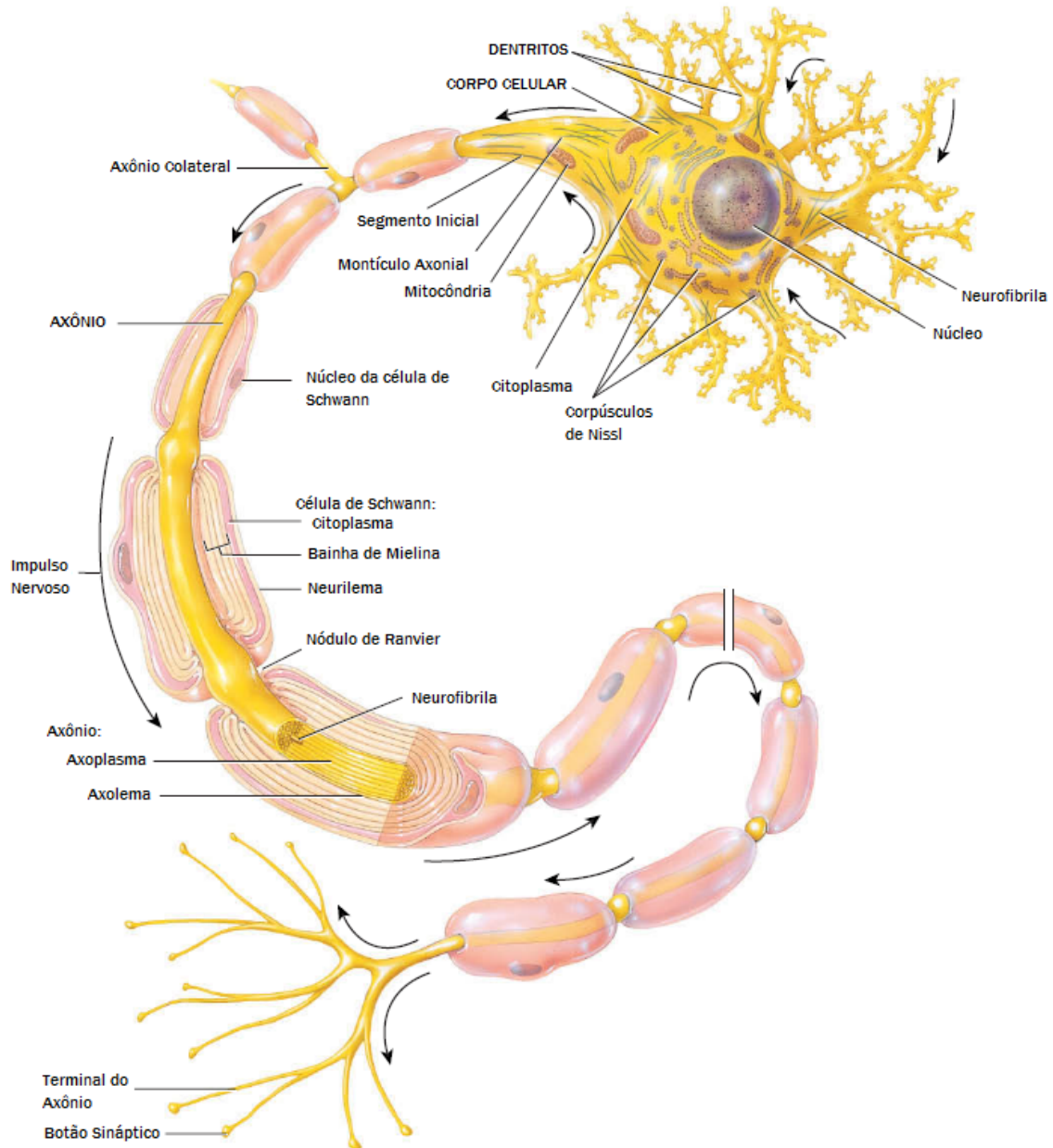
2.1.1 Neurônio

Como as células musculares, neurônios (células nervosas) possuem excitabilidade elétrica - a habilidade de responder a um estímulo e convertê-lo em um potencial de ação. Um estímulo é qualquer mudança no ambiente que é forte o suficiente para iniciar um potencial de ação. Um potencial de ação (impulso nervoso) é um sinal elétrico que se propaga pela superfície da membrana de um neurônio. Ele começa e viaja devido ao movimento dos íons (como sódio e potássio) entre o fluido intersticial e o interior de um neurônio através de canais de íons específicos na sua membrana plasmática. Uma vez iniciado, o impulso nervoso viaja rapidamente (de 0,5 m/s a 130 m/s) a uma intensidade constante (TORTORA; DERRICKSON, 2013).

Tipicamente, os neurônios possuem três partes: o corpo da célula, os dendritos e o axônio. O corpo celular – também conhecido pelos nomes de pericário ou soma – contém o núcleo da célula rodeado pelo citoplasma e organelas (como lisossomos, mitocôndria, complexo de Golgi, ribossomos e retículo endoplasmático). Os dendritos são as partes da célula responsáveis pela recepção de estímulos nervosos e pela retransmissão destes estímulos para o corpo da célula. Eles são, normalmente, curtos e altamente ramificados, de forma a possuírem grandes áreas de contato com outras células. O axônio de um neurônio é a estrutura responsável pela propagação de impulsos nervosos do corpo celular a outros neurônios, às fibras musculares ou a glândulas. Axônios não se ramificam abundantemente como os dendritos. Quando o fazem, dão origem aos chamados axônios colaterais. Em suas extremidades, os axônios e os axônios colaterais dividem-se em múltiplos terminais. Nas pontas dos terminais, encontra-se o botão sináptico, que contém múltiplas vesículas sinápticas que armazenam um produto químico chamado neurotransmissor. Um neurotransmissor é uma substância liberada por uma vesícula sináptica que excita ou inibe outros neurônios, fibras musculares ou glândulas. O botão sináptico, quando estimulado, libera os neurotransmissores, transmitindo, assim, o impulso nervoso. O local de comunicação entre dois neurônios ou entre um neurônio e uma célula atuadora é chamado de

sinapse. A Figura 2.1 mostra a estrutura de um neurônio (TORTORA; DERRICKSON, 2013).

Figura 2.1: A estrutura de um neurônio (adaptado de TORTORA; DERRICKSON, 2013).



O neurônio motor transmite potenciais de ação para fora do sistema nervoso central para atuadores (músculos ou glândulas) na periferia através de nervos cranianos ou espinais (TORTORA; DERRICKSON, 2013).

2.1.2 Potencial de Ação (Impulso Nervoso)

Os sinais nervosos são transmitidos por potenciais de ação, que são rápidas alterações no potencial de membrana que se propagam com grande velocidade pela membrana da fibra nervosa. Cada potencial de ação começa com uma alteração repentina do potencial de repouso normal negativo da membrana para um potencial positivo e, então, termina, com

um retorno quase tão rápido, para o potencial negativo. Para conduzir um sinal nervoso, o potencial de ação viaja pela fibra nervosa até a sua extremidade final (GUYTON; HALL, 2010).

Segundo GUYTON; HALL (2010), os principais estágios sucessivos do potencial de ação são:

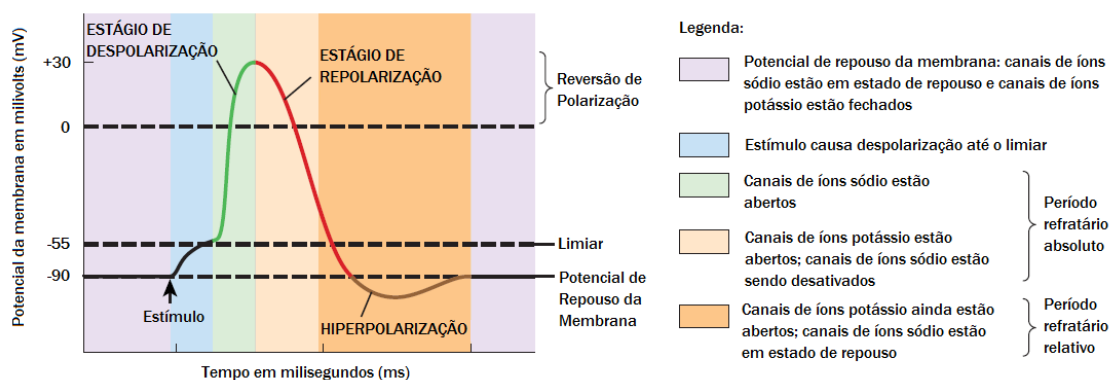
Estágio de Repouso: este é o potencial de repouso da membrana, antes do início do potencial de ação. Diz-se que a membrana está “negativamente polarizada” durante este estágio, uma vez que, no repouso, o potencial de membrana é próximo de -90 milivolts;

Estágio de Despolarização: neste momento, a membrana fica muito permeável aos íons de sódio, permitindo a difusão de uma quantidade muito grande de íons positivamente carregados para o interior do axônio. Com o potencial aumentando para um valor positivo, o estado normal “polarizado” de -90 milivolts é, rapidamente, neutralizado pelo fluxo de íons de sódio. Este processo é chamado de despolarização. Em fibras nervosas largas, o grande excesso de íons positivos que se movem para o interior da fibra faz o potencial de membrana ultrapassar o nível zero e tornar-se positivo. Em algumas fibras menores, assim como em muitos neurônios do sistema central, o potencial meramente se aproxima do nível zero e não o ultrapassa (não chega ao estado positivo);

Estágio de Repolarização: em alguns décimos de milésimos de segundo, após a membrana ter ficado altamente permeável aos íons de sódio, os canais de sódio começam a se fechar, e os canais de potássio se abrem mais que o normal. Então, uma rápida difusão de íons de potássio para o exterior restabelece o potencial de repouso normal negativo da membrana. Este processo é chamado de repolarização da membrana.

A Figura 2.2 ilustra os diferentes estágios do potencial de ação.

Figura 2.2: Potencial de Ação ou Impulso (adaptado de TORTORA; DERRICKSON, 2013).

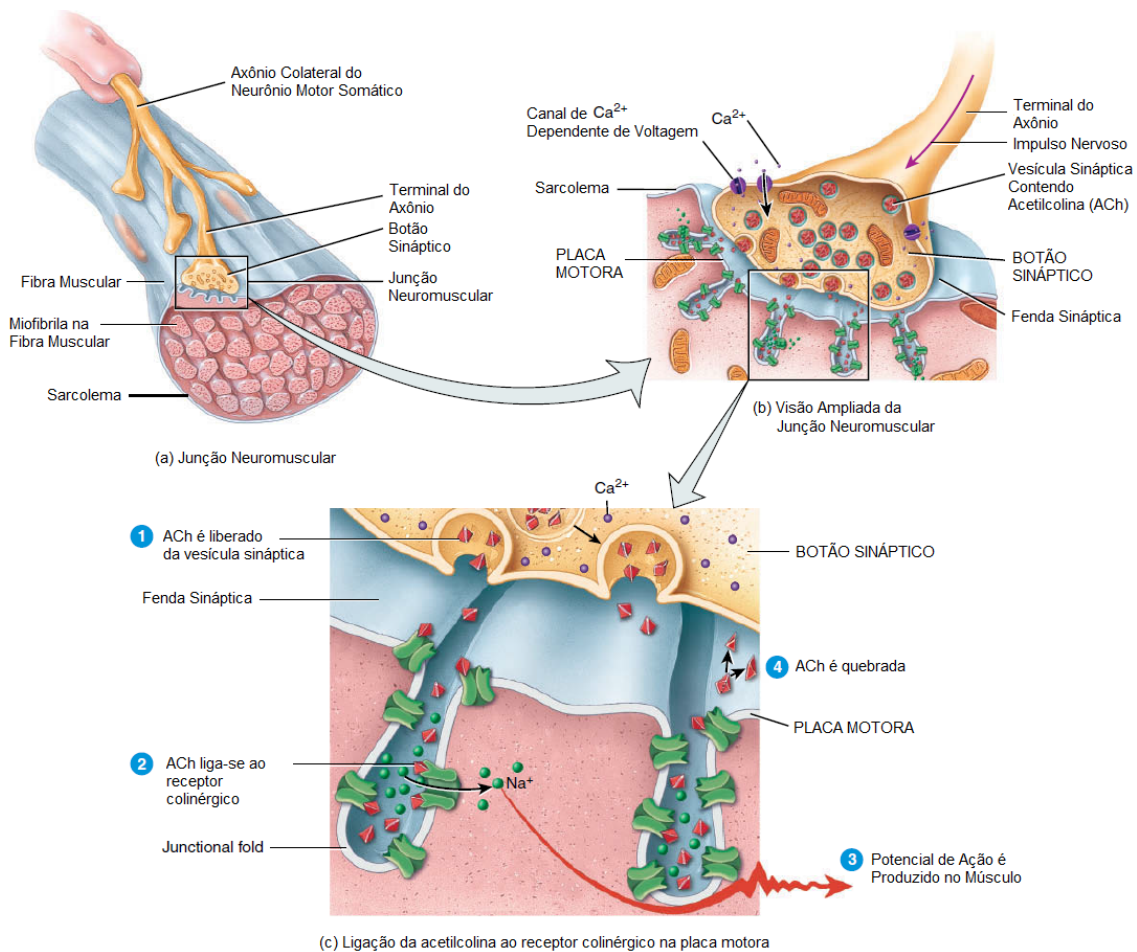


2.1.3 Junção Neuromuscular

Os neurônios que estimulam fibras musculares esqueléticas para a contração são chamados neurônios motores somáticos. Cada neurônio motor somático tem um axônio que

se estende do cérebro ou da coluna vertebral até um grupo de fibras esqueléticas musculares. Uma fibra muscular contrai-se em resposta a um ou mais potenciais de ação se propagando pelo seu sarcolema e pelo seu sistema de túbulos-T. Potenciais de ação musculares aparecem na junção neuromuscular - a sinapse entre um neurônio motor somático e uma fibra esquelética muscular (veja a Figura 2.3). Uma sinapse é a região onde a comunicação ocorre entre dois neurônios, ou entre um neurônio e uma célula-alvo – neste caso, entre um neurônio motor somático e uma fibra muscular. Na maioria das sinapses, uma pequena lacuna, chamada de fenda sináptica, separa as duas células. Como as células não se tocam fisicamente, o potencial de ação não pode “pular a fenda” de uma célula para a outra. Ao invés disso, a primeira célula comunica-se com a segunda através da liberação de mensageiros químicos chamados neurotransmissores. A Figura 2.3 mostra a estrutura da junção neuromuscular (TORTORA; DERRICKSON, 2013).

Figura 2.3: Estrutura da junção neuromuscular, a sinapse entre um neurônio motor somático e uma fibra muscular esquelética (adaptado de TORTORA; DERRICKSON, 2013).



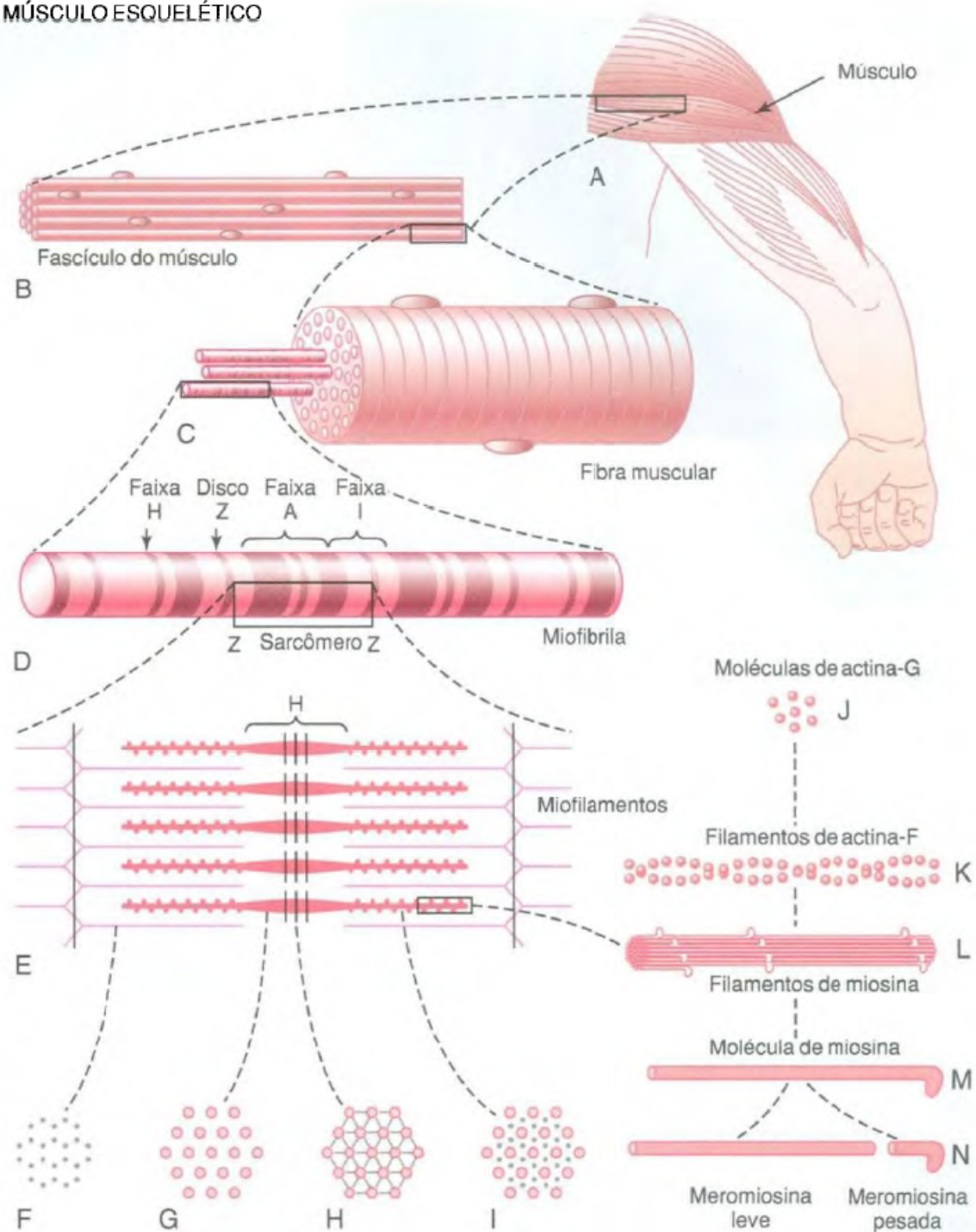
2.1.4 Músculo Esquelético

Segundo BELTRAMINI (1999) e ORTOLAN (2002), o músculo esquelético é composto de fascículos musculares que, por sua vez, são constituídos de células chamadas fibras musculares. As fibras musculares esqueléticas têm a forma de um cilindro com diâmetros variando de $10\mu m$ a $80\mu m$. As fibras, por sua vez, são compostas de centenas

de miofibrilas, de núcleos celulares e do retículo sarcoplasmático. A Figura 2.4 mostra a organização do músculo esquelético.

Figura 2.4: Organização do músculo esquelético, do nível macroscópico ao molecular. As letras F, G, H e I são cortes transversais nos níveis indicados (adaptado de GUYTON; HALL, 2010).

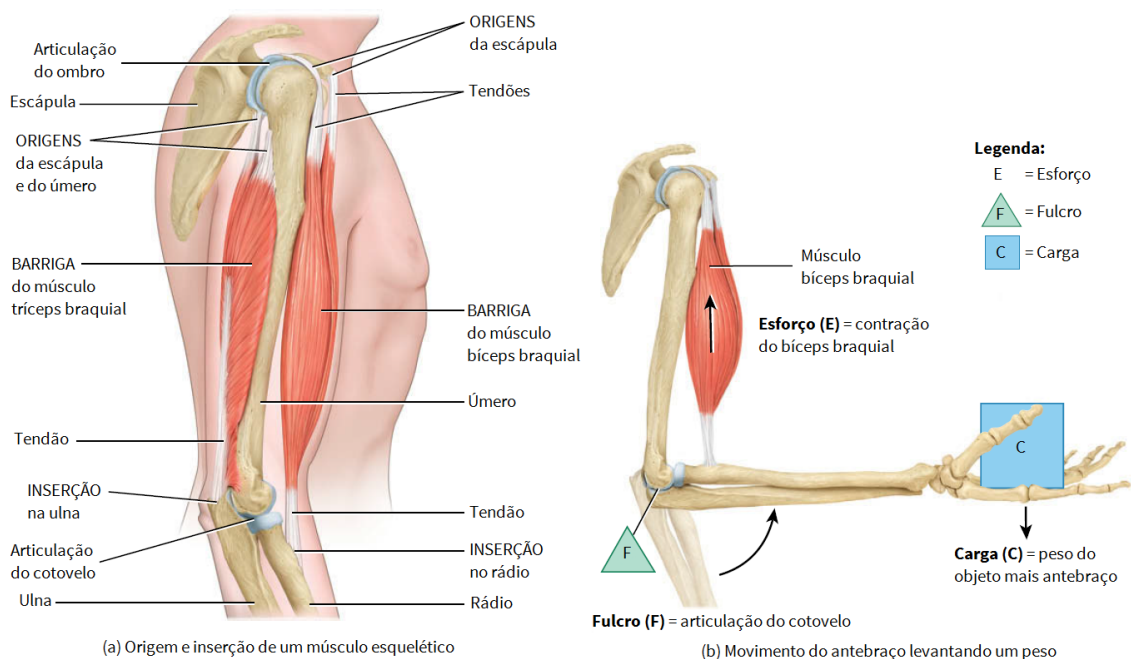
MÚSCULO ESQUELÉTICO



Segundo TORTORA; DERRICKSON (2013), os músculos esqueléticos, que são ligados aos ossos por tendões nas suas origens e inserções, produzem movimentos exercendo

força nos tendões, que, por sua vez, empurram (ou puxam) os ossos ou outras estruturas, como, por exemplo, a pele. A maioria dos músculos cruza ao menos uma articulação e é, normalmente, ligada aos ossos articulados que formam a articulação. Quando um músculo esquelético se contrai, ele move uma das articulações. Os dois ossos articulados geralmente não se movem igualmente em resposta a uma contração. Um osso permanece estacionário ou perto de sua posição original, seja porque outros músculos o estabilizam ao se contraírem e o empurraram na direção oposta, seja porque sua estrutura o torna menos livre para a movimentação. A Figura 2.5 ilustra a relação entre músculos esqueléticos e ossos.

Figura 2.5: Relação entre músculos esqueléticos e ossos. Os ossos servem como alavancas, e as articulações servem como fulcros para as alavancas (adaptado de TORTORA; DERRICKSON, 2013).



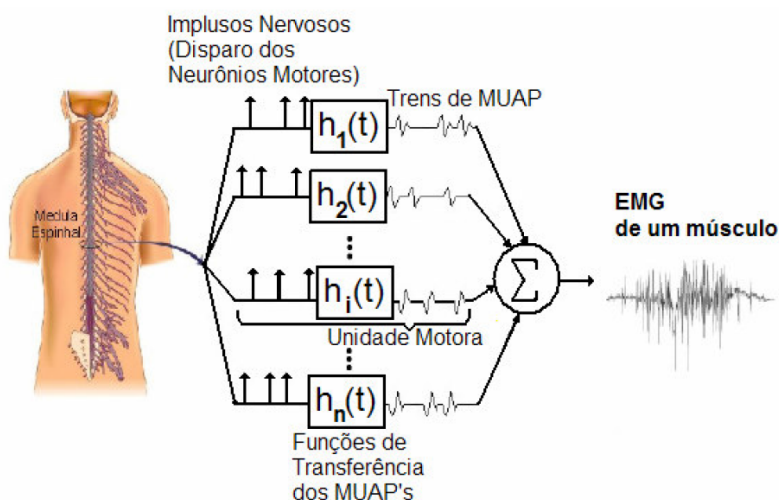
2.2 Eletromiografia

Eletromiografia (EMG) é, de acordo com TORTORA; DERRICKSON (2013), um exame clínico que mede a atividade elétrica (potencial de ação muscular) em músculos relaxados e contraídos através de um equipamento chamado de eletromiógrafo. Normalmente, músculos relaxados produzem muito pouca atividade elétrica; uma contração pequena produz pouca atividade elétrica; e contrações cada vez mais fortes produzem atividade elétrica crescente.

Em eletromiografia, um eletromiógrafo é utilizado para produzir um registro chamado de eletromiograma. Este equipamento detecta o potencial elétrico gerado por células musculares quando estas células são ativas. Os potenciais de ação que percorrem as fibras musculares produzem campos eletromagnéticos em suas proximidades. A resultante destes campos eletromagnéticos pode ser medida com a ajuda de eletrodos próximos ao músculo estudado. Ao registro desses potenciais elétricos gerados nas fibras musculares, dá-se o nome de eletromiografia (FAVIEIRO, 2009, 2012).

O sinal de EMG é a soma dos diversos potenciais de ação de cada músculo, vindos de vários MUAPs (Potencial de Ação de Unidade Motora) de cada unidade motora. A Figura 2.6 mostra uma representação esquemática da geração do sinal mioelétrico a partir do somatório dos trens de MUAPs (RICCIOTTI, 2006).

Figura 2.6: Representação esquemática da geração do sinal mioelétrico (adaptado de RICCIOTTI, 2006).



Hoje, existem dois tipos de eletromiografia em constante uso: eletromiografia intramuscular e eletromiografia superficial. A eletromiografia intramuscular exige a inserção de uma agulha-eletrodo pela pele até atingir o músculo desejado. Desta forma, é possível capturar apenas a atividade de algumas fibras musculares específicas com uma boa intensidade de sinal. Este tipo de eletromiografia é, no entanto, muitas vezes, considerado muito invasivo e desnecessário. A eletromiografia superficial utiliza eletrodos superficiais (sobre a pele) para monitorar a aparência dos sinais de ativação muscular. Ao contrário da eletromiografia intramuscular, esta técnica captura sinais de áreas maiores que apenas algumas fibras musculares, podendo, portanto, sofrer diversas interferências indesejadas – como o *cross-talk* muscular, onde diferentes músculos influenciam em uma mesma captura de sinal. Além disso, os sinais capturados por esta técnica tendem a ser menos intensos. Apesar de menos precisa, a eletromiografia superficial não é invasiva, o que a torna mais apta a utilização em sistemas de tecnologia assistiva (ORTOLAN, 2002; FAVIEIRO, 2009, 2012).

Segundo MAZURKIEWICZ (2007), sinais capturados com eletrodos de superfície possuem muito mais ruído e picos relacionados ao maior número de unidades motoras (MUs). Em sinais como estes, é muito difícil distinguir entre as MUs, especialmente quando a força gerada pelas fibras musculares aumenta. Conforme ORTOLAN (2002), o sinal de EMG captado na superfície da pele possui as seguintes características:

- componentes de frequência entre 0 Hz e 1 kHz, sendo que a maior concentração de energia está dentro da faixa de 50 Hz a 150 Hz;
- amplitudes máximas entre $50\mu V$ e $5mV$.

Ainda segundo ORTOLAN (2002), a amplitude de um sinal detectado superficialmente pode variar de acordo com o tipo de músculo analisado, o nível de contração muscular e o tipo e localização dos eletrodos.

2.3 Musculatura do Antebraço

Segundo BASMAJIAN; DE LUCA (1985), para que o sinal muscular seja capturado da melhor forma possível, é necessário que os eletrodos estejam posicionados sobre os músculos responsáveis pelos movimentos que se deseja analisar. Por isso, é importante relacionar os músculos do corpo humano com os movimentos pelos quais eles são responsáveis. No caso deste projeto, apenas os músculos do antebraço são necessários. A Tabela 2.1 e a Tabela 2.2 relacionam os músculos do antebraço (veja as Figuras 2.7 e 2.8) responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos e suas funções específicas.

Tabela 2.1: Músculos do compartimento anterior do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos e suas funções específicas (adaptado de TORTORA; DERRICKSON, 2013).

Músculo	Ação
Compartimento Anterior do Antebraço Superficial	
Flexor Radial do Carpo <i>(Flexor Carpi Radialis)</i>	Flexão da mão (desvio radial) na articulação do pulso; Abdução da mão (desvio radial) na articulação do pulso.
Palmar Longo <i>(Palmaris Longus)</i>	Flexão fraca da mão na articulação do pulso.
Flexor Ulnar do Carpo <i>(Flexor Carpi Ulnaris)</i>	Flexão da mão (desvio ulnar) na articulação do pulso.
Flexor Superficial dos Dedos <i>(Flexor Digitorum Superficialis)</i>	Flexão da falange média de cada dedo na articulação interfalangeana proximal; Flexão da falange proximal de cada dedo na articulação metacarpofalangeana; Flexão da mão na articulação do pulso.
Compartimento Anterior do Antebraço Profundo	
Flexor Longo do Polegar <i>(Flexor Pollicis Longus)</i>	Flexão da falange distal do polegar na articulação interfalangeana.
Flexor Profundo dos Dedos <i>(Flexor Digitorum Profundus)</i>	Flexão da falange distal de cada dedo na articulação interfalangeana; Flexão da falange média de cada dedo na articulação interfalangeana; Flexão da falange proximal de cada dedo na articulação metacarpofalangeana; Flexão da mão na articulação do punho.

Figura 2.7: Músculos do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos. Da esquerda para a direita, visões anterior superficial, anterior intermediária e anterior profunda (adaptado de TORTORA; DERRICKSON, 2013).

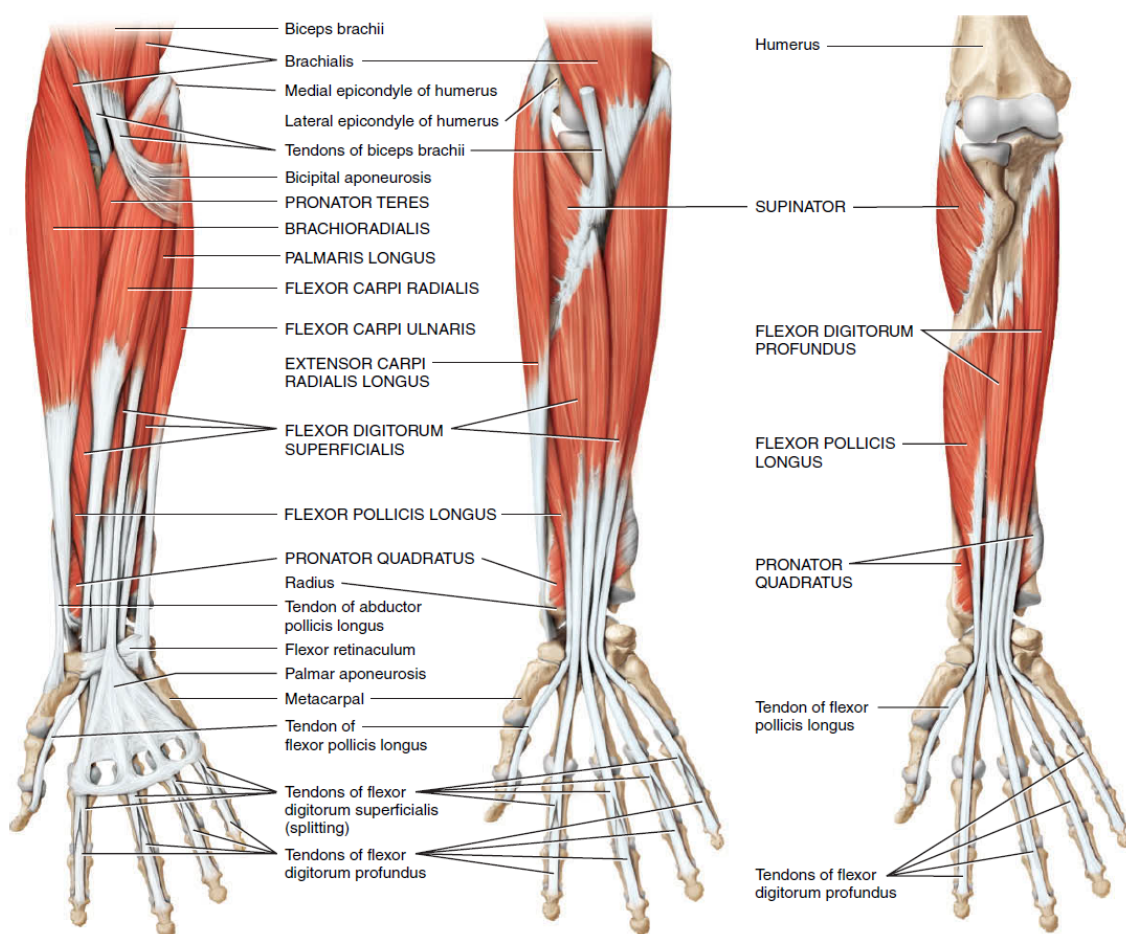
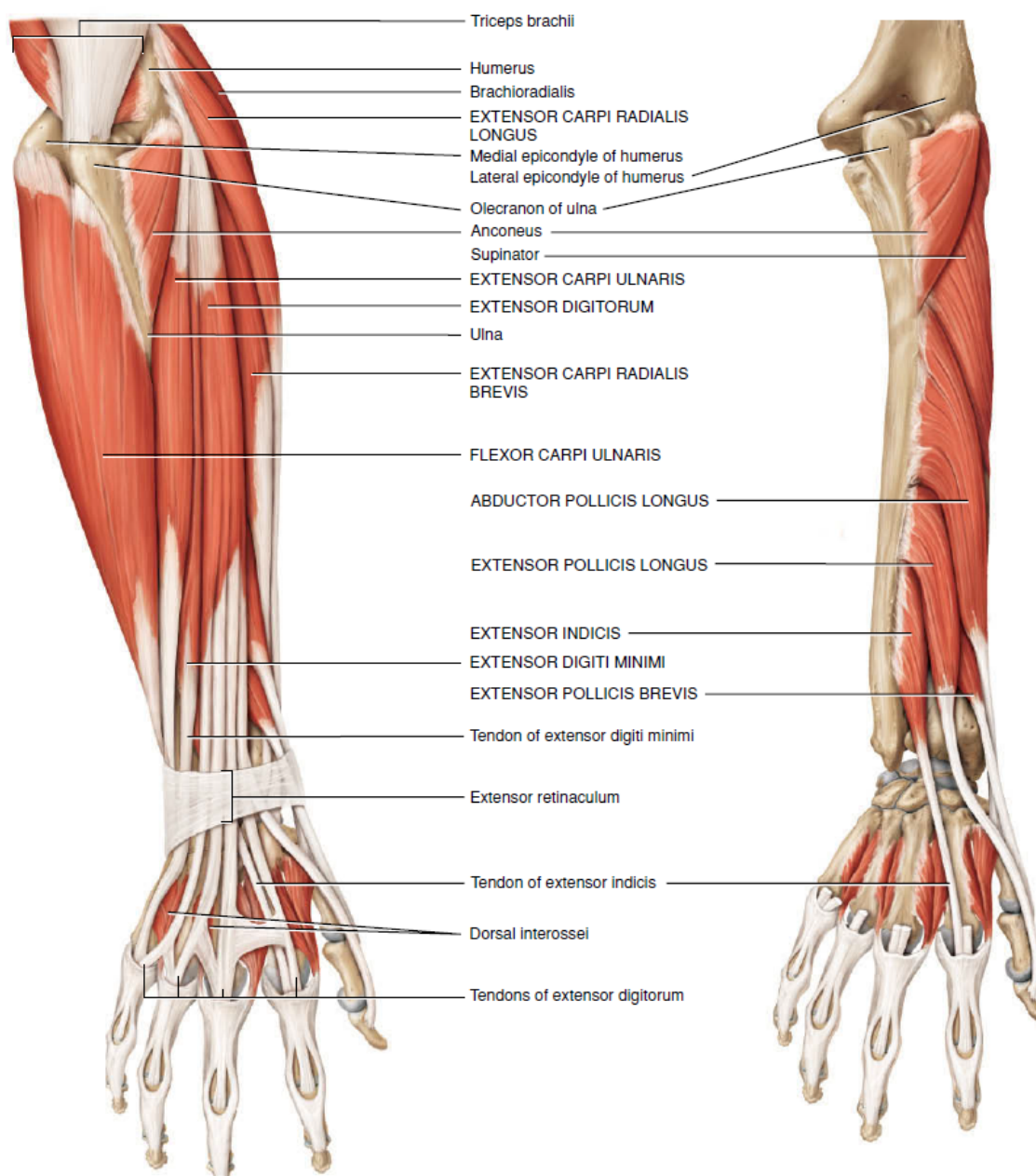


Tabela 2.2: Músculos do compartimento posterior do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos e suas funções específicas (adaptado de TORTORA; DERRICKSON, 2013).

Músculo	Ação
Compartimento Posterior do Antebraço Superficial	
Extensor Radial Longo do Carpo <i>(Extensor Carpi Radialis Longus)</i>	Extensão da mão (desvio ulnar) na articulação do pulso; Abdução da mão (desvio ulnar) na articulação do pulso.
Extensor Radial Curto do Carpo <i>(Extensor Carpi Radialis Brevis)</i>	Extensão da mão na articulação do pulso; Abdução da mão na articulação do pulso.
Extensor dos Dedos <i>(Extensor Digitorum)</i>	Extensão da falange distal de cada dedo na articulação interfalangeana; Extensão da falange média de cada dedo na articulação interfalangeana; Extensão da falange proximal de cada dedo na articulação metacarpofalangeana; Extensão da mão na articulação do pulso.
Extensor do Dedo Mínimo <i>(Extensor Digiti Minimi)</i>	Extensão da falange proximal do dedo mínimo na articulação metacarpofalangeana; Extensão da mão na articulação do pulso.
Extensor Ulnar do Carpo <i>(Extensor Carpi Ulnaris)</i>	Extensão da mão (desvio ulnar) na articulação do pulso; Abdução da mão (desvio ulnar) na articulação do pulso.
Compartimento Posterior do Antebraço Profundo	
Abdutor Longo do Polegar <i>(Abductor Pollicis Longus)</i>	Abdução do polegar na articulação carpometacarpal; Extensão do polegar na articulação carpometacarpal; Abdução da mão na articulação do pulso.
Extensor Curto do Polegar <i>(Extensor Pollicis Brevis)</i>	Extensão da falange proximal do polegar na articulação metacarpofalangeana; Extensão do primeiro metacarpal do polegar na articulação carpometacarpal; Extensão da mão na articulação do pulso.
Extensor Longo do Polegar <i>(Extensor Pollicis Longus)</i>	Extensão da falange distal do polegar na articulação interfalangeana; Extensão do primeiro metacarpal do polegar na articulação carpometacarpal; Abdução da mão na articulação do pulso.
Extensor do Indicador <i>(Extensor Indicis)</i>	Extensão da falange distal do dedo indicador na articulação interfalangeana; Extensão da falange média do dedo indicador na articulação interfalangeana; Extensão da falange proximal do dedo indicador na articulação metacarpofalangeana; Extensão da mão na articulação do pulso.

Figura 2.8: Músculos do antebraço responsáveis pela movimentação do pulso, da mão, do polegar e dos outros dedos. Da esquerda para a direita, visões posterior superficial e posterior profunda (adaptado de TORTORA; DERRICKSON, 2013).



3 BASE DE DADOS NINAPRO

O projeto NinaPro (*Non-Invasive Adaptive Prosthetics*) é um trabalho em curso que visa ajudar, através da criação de bases de dados públicas, a pesquisa sobre o controle de próteses mioelétricas da mão. Os bancos de dados são obtidos através do registro de sinais de eletromiografia superficial e de dados cinemáticos da mão enquanto os sujeitos em estudo realizam um conjunto de ações e poses predefinidas (GIJSBERTS *et al.*, 2014).

Este trabalho utiliza a segunda base de dados NinaPro, que contém dados coletados de 40 pessoas não amputadas para um total de 49 movimentos distintos (além do estado de repouso) (GIJSBERTS *et al.*, 2014).

3.1 Protocolo de Aquisição de Dados

O experimento - do qual são coletados os sinais mioelétricos contidos na base de dados NinaPro - é dividido em três exercícios: movimentos básicos dos dedos e do pulso (poses 1 a 17 da Figura 3.2), movimentos de pegada e funcionais (poses 18 a 40 da Figura 3.2) e padrões de força (poses 41 a 49 da Figura 3.2). Nos primeiros dois exercícios, os sujeitos devem repetir uma série de movimentos representados por filmes que são mostrados em uma tela. No último exercício, os sujeitos devem pressionar combinações de dedos com uma força crescente (GIJSBERTS *et al.*, 2014).

Conforme GIJSBERTS *et al.* (2014), a atividade muscular é adquirida através de doze eletrodos ativos sem fio (o *Delsys™ Trigno Wireless System*[®]). Oito eletrodos são posicionados, de forma igualmente espaçada, em volta do antebraço, em correspondência com a articulação radioumeral; dois eletrodos são posicionados nos pontos de maior atividade dos músculos *flexor digitorum superficialis* (flexor superficial dos dedos) e *extensor digitorum communis* (extensor dos dedos); os outros dois eletrodos são posicionados nos pontos de maior atividade do bíceps e do tríceps. Estas localizações foram escolhidas de forma a combinar uma amostragem densa com uma estratégia de posicionamento anatômico precisa. Os sinais mioelétricos são amostrados a uma taxa de 2 kHz com um ruído menor que 750 nV RMS. A Figura 3.1 mostra o posicionamento dos doze eletrodos.

A informação cinemática é registrada com a ajuda de uma luva de aquisição de dados (a *CyberGlove II*), de um inclinômetro (localizado no pulso do sujeito) e de acelerômetros (integrados em cada um dos eletrodos de superfície). Os dados destes acelerômetros de três eixos são amostrados a uma taxa de 148 Hz (GIJSBERTS *et al.*, 2014).

A intensidade de força é medida com a ajuda de um sensor de força que foi especialmente desenvolvido para o experimento chamado *Finger-Force Linear Sensor* (FFLS). Este dispositivo mede forças de flexão e de extensão dos dedos 2-5 (dedos indicador, médio, anular e mínimo), usando um extensômetro uniaxial linear, assim como as forças de flexão, de extensão, de abdução e de adução do dedo 1 (polegar), usando um sensor

Figura 3.1: Posicionamento dos doze eletrodos superficiais no braço. O eletrodo do músculo flexor superficial dos dedos está oculto pelo braço e, portanto, não é visível na imagem (adaptado de GIJSBERTS *et al.*, 2014).



similar de eixo duplo. Estes sensores são caracterizados por sinais de alta repetibilidade, *drift* mínimo com o tempo, linearidade quase perfeita e histerese virtualmente inexistente (parâmetros desviam não mais que 0,3%). Cada sensor de força é ligado a um amplificador dedicado, cuja saída é amostrada a uma taxa de 100 Hz por uma placa de aquisição de dados (DAQ) da National InstrumentsTM - o NI-DAQ PCMCIA 6024E, de resolução de 12 bits (GIJSBERTS *et al.*, 2014).

De um modo geral, o processo de aquisição de dados funciona da seguinte maneira. Primeiramente, o computador exibe, na tela de estímulo, um modelo virtual realizando um dos movimentos considerados. O sujeito em teste, que já se encontra equipado - com eletrodos de superfície, acelerômetros e luva de aquisição de dados cinemáticos e inclinômetro (no caso dos dois primeiros exercícios) ou *Finger-Force Linear Sensor* (no caso do último exercício) -, repete o movimento executado pelo modelo virtual. Os sinais de atividade muscular - produzidos pelo sujeito em teste ao repetir o movimento mostrado na tela de estímulo - são captados pelos eletrodos de superfície, condicionados e convertidos pela base do equipamento de eletromiografia e repassados ao computador - que os registra de forma permanente. Os dados cinemáticos são adquiridos pela luva de aquisição de dados, pelo inclinômetro e pelos acelerômetros (contidos nos eletrodos) e são, também, repassados e armazenados pelo computador. Os dados de intensidade de força são capturados pelos sensores de força (FFLS) e são, igualmente, armazenados pelo computador. Estes são os sinais contidos na base de dados NinaPro (GIJSBERTS *et al.*, 2014).

Durante a aquisição, pede-se que os sujeitos repitam os movimentos com a mão direita. Cada repetição de movimento dura 5 segundos e é seguida por 3 segundos de repouso. O protocolo inclui 6 repetições de 49 movimentos diferentes realizados por 40 sujeitos intactos. A Figura 3.2 mostra os movimentos considerados pela base de dados NinaPro (GIJSBERTS *et al.*, 2014).

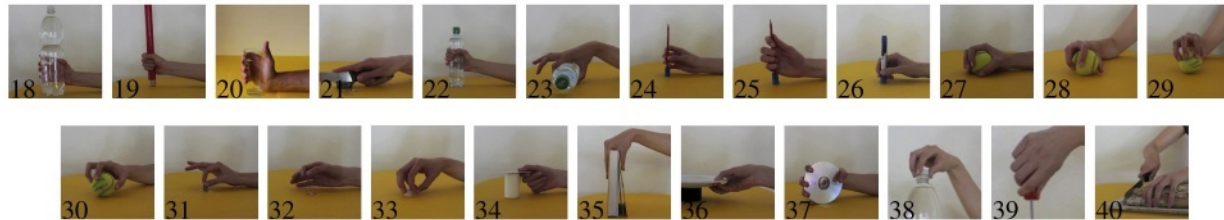
Figura 3.2: Os 49 movimentos (ou posturas) mais posição de repouso do segundo banco de dados NinaPro (adaptado de GIJSBERTS *et al.*, 2014).



(a) Configurações isométrica e isotônica da mão.



(b) Movimentos básicos do pulso.



(c) Movimentos de pegada e funcionais.



(d) Medida de Padrões de Força dos Dedos.

(e) Posição de Repouso.



3.2 Conjunto de Dados

A base de dados NinaPro armazena os sinais captados em um conjunto de arquivos MAT (formato de arquivo interno do MathWorks® MATLAB®). Para cada um dos sujeitos contidos na base de dados, existem três arquivos MAT - um para cada um dos exercícios do experimento.

Os arquivos MAT incluem as seguintes variáveis:

- ***subject***: número de identificação do sujeito;
- ***exercise***: número de identificação do exercício;
- ***emg***: matriz de N linhas e 12 colunas, onde cada linha representa uma amostra no tempo, e cada coluna representa o sinal captado por um dos 12 eletrodos de superfície;
- ***acc***: matriz de N linhas e 36 colunas, onde cada linha representa uma amostra no tempo, e cada coluna representa o sinal de um dos 3 eixos de um dos 12 acelerômetros;
- ***glove***: matriz de N linhas e 22 colunas, onde cada linha representa uma amostra no tempo, e cada coluna representa o sinal de um dos 22 sensores de inclinação da luva de aquisição de dados;
- ***inclin***: matriz de N linhas e 2 colunas, onde cada linha representa uma amostra no tempo, e cada coluna representa o sinal de um dos 2 eixos do inclinômetro do pulso do sujeito;
- ***stimulus***: matriz de N linhas e 1 coluna, onde cada linha representa uma amostra no tempo do número de identificação do movimento sendo executado naquela amostra;
- ***repetition***: matriz de N linhas e 1 coluna, onde cada linha representa uma amostra no tempo do número de identificação da repetição do movimento sendo executado naquela amostra;
- ***restimulus***: matriz *stimulus* refinada em pós-processamento;
- ***rerepetition***: matriz *repetition* refinada em pós-processamento;
- ***force***: matriz de N linhas e 6 colunas, onde cada linha representa uma amostra no tempo, e cada coluna representa o sinal de um dos 6 sensores de força do FFLS;
- ***forcecal***: matriz de 2 linhas e 6 colunas, onde cada coluna representa o valor de calibração de um dos 6 sensores de força do FFLS, a primeira linha corresponde aos valores mínimos, e a segunda linha corresponde aos valores máximos.

As variáveis *subject*, *exercise*, *acc*, *emg*, *stimulus* e *repetition* estão contidas em todos os arquivos MAT; as variáveis *glove*, *inclin*, *rerepetition* e *restimulus* estão presentes somente nos arquivos correspondentes aos exercícios 1 e 2; e as variáveis *force* e *forcecal* fazem parte apenas dos arquivos que representam o exercício 3.

Este trabalho utiliza apenas os dados contidos nas variáveis *subject*, *exercise*, *emg*, *stimulus* e *repetition*.

3.3 Conjunto de Poses Seleccionadas

Entre o conjunto de posturas contidas no banco de dados NinaPro (veja a Figura 3.2), as poses que serão utilizadas neste trabalho são: abdução dos dedos (imagem 5), dedos flexionados em punho (imagem 6), extração do dedo indicador (imagem 7), flexão da mão (imagem 13), extensão da mão (imagem 14) e pegada (imagem 29). A Figura 3.3 exibe o conjunto de movimentos considerados neste trabalho.

Figura 3.3: Os 6 movimentos (ou posturas) seleccionados do segundo banco de dados NinaPro (adaptado de GIJSBERTS *et al.*, 2014).



4 PRÉ-PROCESSAMENTO

A etapa de pré-processamento consiste em uma série de tratamentos do sinal muscular antes que ele possa ser efetivamente utilizado por um algoritmo de aprendizado (estes algoritmos são descritos no Capítulo 5). Esta etapa é dividida em: ajuste de *offset*, filtragem, normalização, segmentação e caracterização (extração de características).

4.1 Ajuste de *Offset*

Esta etapa é responsável por remover o *offset* do sinal, ou tornar a média do sinal (como um todo) igual a zero. Para isto, atualiza-se o vetor que contém o sinal considerado através da atribuição representada pela Equação (4.1):

$$EMG = EMG - \frac{1}{N} \sum_{i=1}^N EMG_{i1}, \quad (4.1)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, e EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz.

A fim ilustrativo, a Figura 4.1 mostra um sinal da forma

$$y = 1 + 2 \sin(2\pi x) \quad (4.2)$$

antes e após a etapa de ajuste de *offset*. Observe que, antes do ajuste, a senoide possui um deslocamento vertical de 1 unidade e, após o ajuste, o deslocamento vertical deixa de existir.

4.2 Filtragem

Esta etapa tem o intuito de eliminar frequências consideradas menos relevantes de um sinal. A filtragem do sinal é realizada com um filtro digital passa-banda com frequências de corte 20 Hz e 500 Hz – faixa de frequência de maior energia de um sinal mioelétrico do segmento mão-braço –, de quarta ordem e com resposta Butterworth, como sugerido por RICCIOTTI (2006) e NILSON (2014). A função de transferência no domínio Z deste filtro é representada pela Equação (4.3):

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{i=1}^{2n+1} b_i z^{i-1}}{\sum_{i=1}^{2n+1} a_i z^{i-1}}, \quad (4.3)$$

onde $a = [1,0000; -3,9029; 6,2959; -5,9129; 4,0584; -2,1089; 0,6513; -0,0990; 0,0181]$, $b = [0,0827; 0,0000; -0,3307; 0,0000; 0,4960; 0,0000; -0,3307; 0,0000; 0,0827]$, e $n = 4$ é a ordem do filtro. A Figura 4.2 mostra a resposta em frequência deste filtro.

Figura 4.1: Exemplo de ajuste de *offset* do sinal representado pela Equação (4.2).

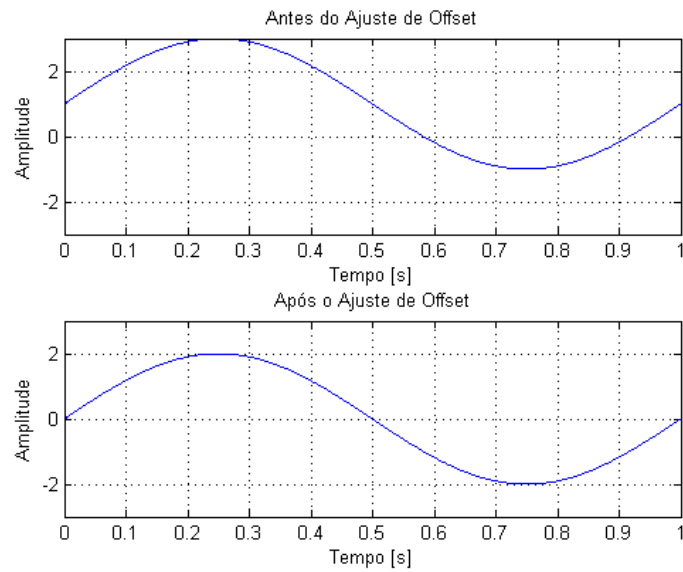
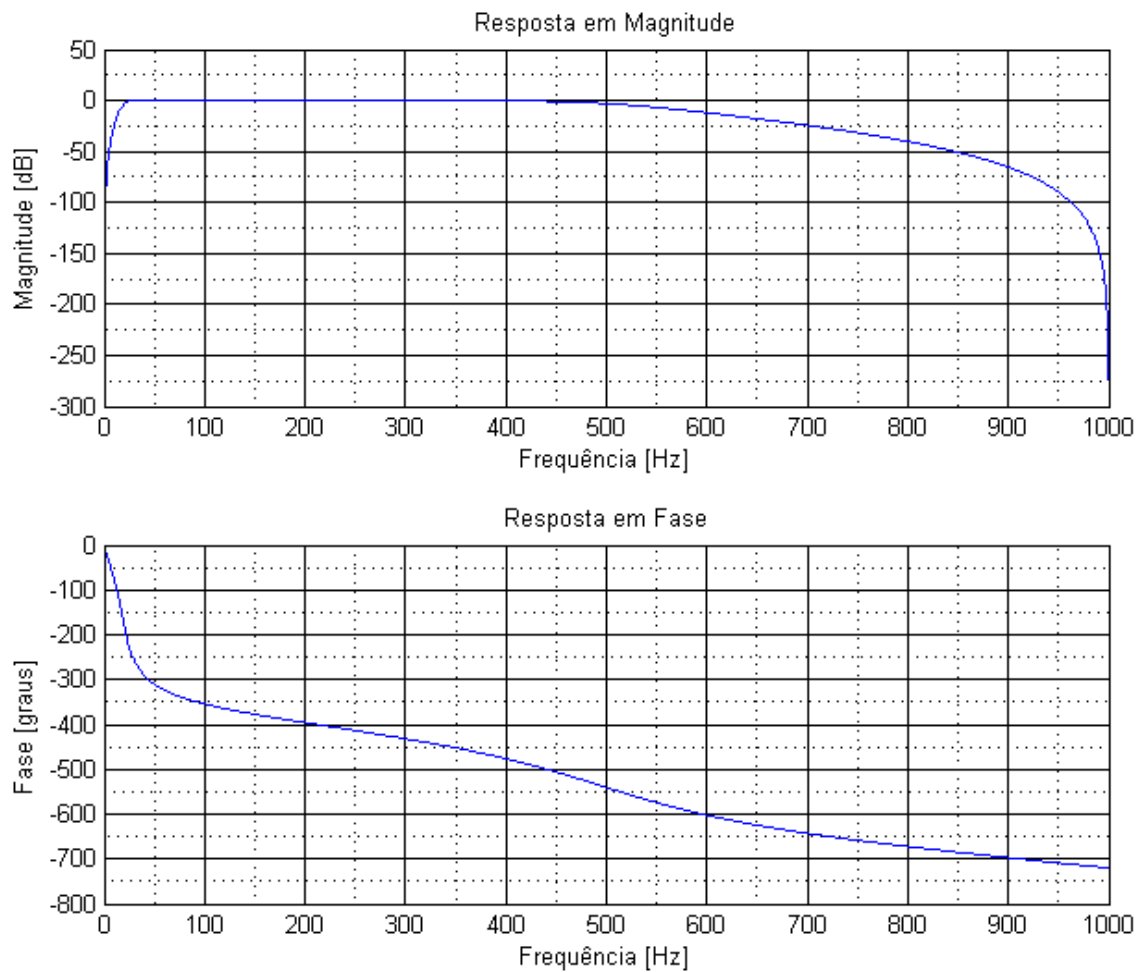


Figura 4.2: Resposta em frequência do filtro digital passa-banda com frequências de corte 20 Hz e 500 Hz de quarta ordem e com resposta Butterworth.



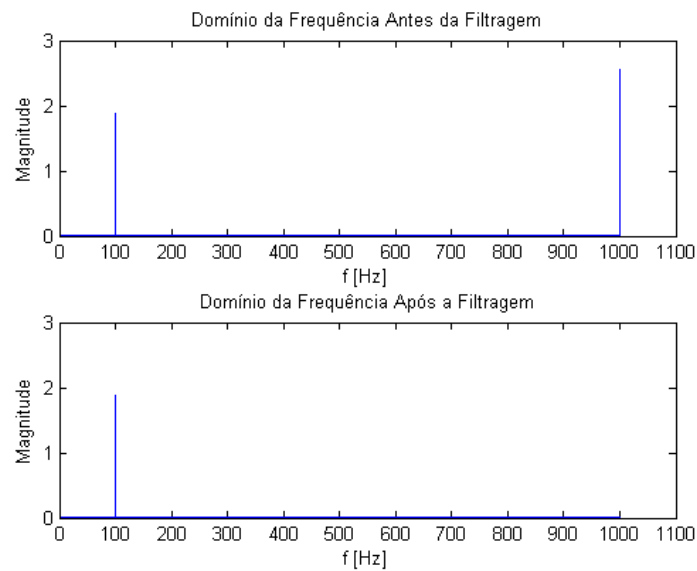
O filtro Butterworth é desenvolvido de forma a obter a resposta em frequência mais plana possível na banda passante e a se aproximar de zero na banda rejeitada. Apesar de ter um *roll-off* mais lento que filtros Chebyshev e elípticos, o filtro Butterworth tem uma fase de resposta mais linear na banda passante (NILSON, 2014).

A Figura 4.3 mostra, como exemplo, o espectro de frequência de um sinal da forma

$$y = 2 \sin(2\pi 100x) + 2 \sin(2\pi 1000x) \quad (4.4)$$

antes e após a etapa de filtragem. Observe que, antes da filtragem, existem dois picos de frequência no espectro (100 Hz e 1000 Hz), e, após a filtragem, o pico de frequência em 1000 Hz deixa de existir.

Figura 4.3: Espectro de frequência do sinal representado pela Equação (4.4) antes e após a filtragem.



4.3 Normalização

A etapa de normalização consiste em mapear o sinal a uma amplitude em $[-1, +1]$. Para isto, atualiza-se o vetor que contém o sinal considerado através da atribuição representada pela Equação (4.5):

$$EMG = \frac{EMG}{\max(|EMG|)}, \quad (4.5)$$

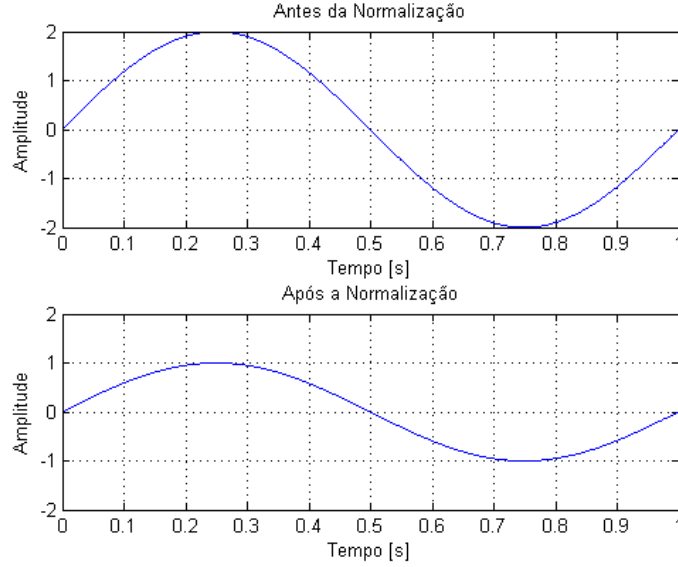
onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, e a operação $\max(|EMG|)$ recupera o elemento de valor máximo entre o valor absoluto dos elementos deste vetor.

A fim ilustrativo, a Figura 4.4 mostra um sinal da forma

$$y = 2 \sin(2\pi x) \quad (4.6)$$

antes e após a etapa de normalização. Observe que, antes da normalização, a senoide possui uma amplitude de 2 unidades e, após a normalização, a amplitude torna-se 1 unidade.

Figura 4.4: Exemplo de normalização do sinal representado pela Equação (4.6).



4.4 Segmentação

A etapa de segmentação tem o intuito de identificar segmentos do sinal que apresentem atividade muscular. O procedimento de segmentação consiste, primeiramente, em dividir o sinal em janelas de 250 ms. Em seguida, cada uma das janelas é classificada. Caso a potência média do sinal da janela

$$P = \frac{1}{N_{win}} \sum_{i=1}^{N_{win}} EMGwin_{i1}^2, \quad (4.7)$$

onde $EMGwin$ é uma matriz $N_{win} \times 1$ que representa as N_{win} amostras de sinal da janela considerada e $EMGwin_{i1}$ é o elemento de linha i e de coluna 1 desta matriz, seja maior que um limiar especificado, a janela é classificada como válida; caso o contrário, ela é classificada como inválida. O valor limiar mencionado depende do valor absoluto máximo do sinal, $\max(|EMG|)$. Caso a intensidade máxima do sinal, $\max(|EMG|)$, seja maior ou igual a 30 vezes a média do valor absoluto do sinal

$$\overline{|EMG|} = (1/N) \sum_{i=1}^N |EMG_{i1}|, \quad (4.8)$$

então o limiar é definido como

$$threshold = 4\% \text{ de } 5\overline{|EMG|}; \quad (4.9)$$

caso o contrário, o limiar é

$$threshold = 4\% \text{ de } \frac{\max(|EMG|)}{5}, \quad (4.10)$$

ou, mais compactamente,

$$threshold = 4\% \text{ de } \begin{cases} 5\overline{|EMG|} & \text{se } \max(|EMG|) \geq 30\overline{|EMG|} \\ \max(|EMG|)/5 & \text{se } \max(|EMG|) < 30\overline{|EMG|} \end{cases},$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras do sinal considerado, EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz, $\overline{|EMG|}$ é a média do valor absoluto do sinal, e a operação $\max(|EMG|)$ recupera o elemento de valor máximo entre o valor absoluto dos elementos deste vetor. O método de segmentação descrito aqui foi sugerido por NILSON (2014).

A Figura 4.5 mostra, como exemplo, um sinal da forma

$$y = y_1 + y_2, \quad (4.11)$$

onde

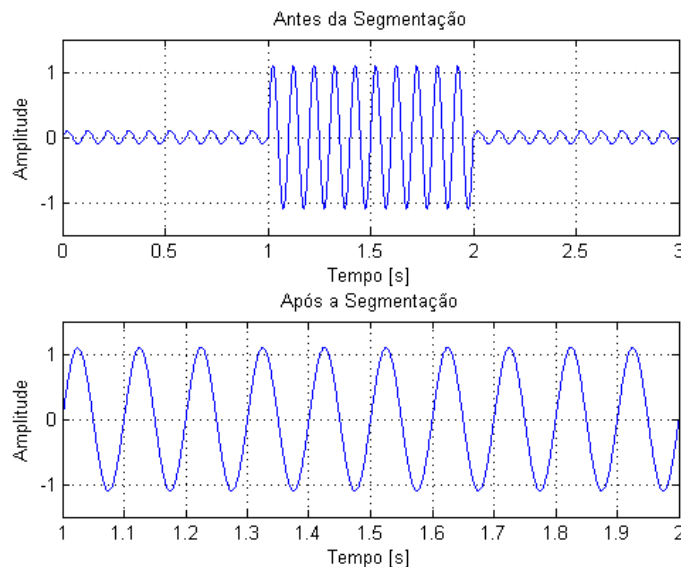
$$y_1 = 0.1 \sin(2\pi 10x_1) \quad x_1 \in [0, 3] \quad (4.12)$$

e

$$y_2 = \sin(2\pi 10x_2) \quad x_2 \in [1, 2], \quad (4.13)$$

antes e após a etapa de segmentação. Observe que, antes da segmentação, o sinal possui áreas de sinal menos intenso ($x \in [0, 1] \cup [2, 3]$) e áreas de sinal mais intenso ($x \in [1, 2]$) e, após a segmentação, somente a área de sinal mais intenso ($x \in [1, 2]$) continua a existir.

Figura 4.5: Sinal representado pela Equação (4.11) antes e após a segmentação.



4.5 Extração de Características

A etapa de caracterização consiste em extrair características do sinal que o descrevam de forma mais concisa, diminuindo a quantidade de dados a serem manipulados e aumentando a quantidade de informação que esses dados carregam sobre a onda. Estas características serão utilizadas como entrada para os métodos de aprendizado de máquina, que serão descritos no Capítulo 5.

A extração de características é realizada nos domínios do tempo e da frequência. Para a caracterização dos sinais no domínio do tempo e no domínio da frequência, são utilizadas as seguintes medidas - que foram sugeridas por NILSON (2014): RMS, média aritmética, variância, desvio padrão, obliquidade e curtose. Além dessas, uma outra característica é identificada no espectro de frequência: o centroide espectral.

4.5.1 RMS

O valor quadrático médio (ou RMS) é uma medida estatística que representa a magnitude de uma quantidade variável. O cálculo do valor RMS do sinal é realizado conforme a Equação (4.14):

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N (EMG_{i1})^2}, \quad (4.14)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, e EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz.

4.5.2 Média Aritmética

A média aritmética representa a medida de tendência central de uma distribuição. Ela pode ser calculada conforme a Equação (4.15):

$$\overline{EMG} = \frac{1}{N} \sum_{i=1}^N EMG_{i1}, \quad (4.15)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, e EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz.

4.5.3 Variância

A variância quantifica a dispersão dos pontos de uma distribuição em comparação com a média. Ela pode ser calculada conforme a Equação (4.16):

$$s^2 = \frac{1}{N-1} \sum_{i=1}^N (EMG_{i1} - \overline{EMG})^2, \quad (4.16)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz, e \overline{EMG} é a média aritmética do sinal (como especificado na Seção 4.5.2).

4.5.4 Desvio Padrão

O desvio padrão, que está diretamente relacionado à variância, quantifica a dispersão dos pontos de uma distribuição em comparação com a média. Ele pode ser calculado conforme a Equação (4.17):

$$s = \sqrt{s^2}, \quad (4.17)$$

onde s^2 é a variância do sinal (como especificado na Seção 4.5.3).

4.5.5 Obliquidade

Obliquidade é uma medida de assimetria de uma distribuição em relação à média. Ela pode ser calculada conforme a Equação (4.18):

$$Obliquidade = \frac{\frac{1}{N} \sum_{i=1}^N (EMG_{i1} - \overline{EMG})^3}{\left(\sqrt{\frac{1}{N} \sum_{i=1}^N (EMG_{i1} - \overline{EMG})^2} \right)^3}, \quad (4.18)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz, e \overline{EMG} é a média aritmética do sinal (como especificado na Seção 4.5.2).

4.5.6 Curtose

A curtose é uma medida de dispersão que caracteriza o pico ou o “achatamento” da distribuição. Ela pode ser calculada conforme a Equação (4.19):

$$Curtose = \frac{\frac{1}{N} \sum_{i=1}^N (EMG_{i1} - \overline{EMG})^4}{\left(\frac{1}{N} \sum_{i=1}^N (EMG_{i1} - \overline{EMG})^2\right)^2}, \quad (4.19)$$

onde EMG é uma matriz $N \times 1$ que representa as N amostras de sinal consideradas, EMG_{i1} é o elemento de linha i e de coluna 1 desta matriz, e \overline{EMG} é a média aritmética do sinal (como especificado na Seção 4.5.2).

4.5.7 Centroide Espectral

O centroide espectral indica o baricentro (ou o “centro de massa”) do espectro de frequência de um sinal. Ele é calculado como a média ponderada das frequências presentes em um sinal, onde os pesos são as magnitudes do espectro. Assim,

$$Centroide = \frac{\sum_{i=1}^N i f_{i1}}{\sum_{i=1}^N f_{i1}}, \quad (4.20)$$

onde f é uma matriz $N \times 1$ que representa a magnitude (ou peso), no espectro de frequência, das N frequências consideradas, e f_{i1} é o elemento de linha i e de coluna 1 desta matriz (PEETERS, 2004).

4.6 Feature Scaling

Feature Scaling é um método de padronização de intervalos de dados. Segundo LIN; HSU; CHANG (2003), a principal vantagem da aplicação deste procedimento é evitar que características cujos possíveis valores estão distribuídos em um grande intervalo dominem outras cujos valores estão distribuídos em intervalos menores. Além disso, ainda de acordo com o autor mencionado, o procedimento evita dificuldades numéricas durante os cálculos.

A normalização dos dados determina a efetiva normalização dos parâmetros (pesos), o que pode exercer grande influência na qualidade do resultado final da aplicação destes dados em um algoritmo de aprendizado. Este procedimento garante que todas as entradas serão tratadas igualmente no processo de regularização (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Como sugerido por HASTIE; TIBSHIRANI; FRIEDMAN (2009), neste trabalho, as características são padronizadas de forma a terem média 0 e desvio padrão 1 (método chamado de padronização). Para isto, atualiza-se o vetor de características através da atribuição representada pela Equação (4.21):

$$features = \frac{features - \overline{features}}{std(features)} \quad (4.21)$$

onde $features$ é uma matriz $N \times 1$ que representa N exemplos de uma das características, $\overline{features}$ é a média aritmética destes N exemplos, e a operação $std(features)$ recupera o desvio padrão deste vetor.

É importante notar que o procedimento de *feature scaling* é aplicado ao fim do pré-processamento de todos os sinais, quando todos os exemplos já foram identificados.

5 APRENDIZADO DE MÁQUINA

De acordo com BISHOP (2006) e com HASTIE; TIBSHIRANI; FRIEDMAN (2009), modelos de aprendizado de máquina são modelos estatísticos capazes de aprender e reconhecer padrões em conjuntos de dados. Tipicamente, em modelos de classificadores, como os utilizados neste trabalho, deseja-se prever a classe (ou categoria) a qual um conjunto de dados de entrada (características) pertence. Para isto, precisa-se de um conjunto de exemplos de treinamento, que deve conter tanto as características quanto a classe a qual cada exemplo pertence. Com estes dados, pode-se construir um modelo de aprendizado capaz de prever a qual categoria um novo conjunto de características (nunca visto antes pelo modelo) pertence. Este tipo de modelo é chamado de modelo de aprendizado supervisionado, pois necessita, para o processo de aprendizado, que os exemplos contenham a classe de saída esperada.

No caso deste trabalho, as classes correspondem ao tipo de movimento representado por um sinal (veja as poses consideradas pela base de dados NinaPro na Figura 3.2), e o conjunto de dados de entrada é o vetor de características extraídas desse sinal (como visto na Seção 4.5). Os modelos de aprendizado utilizados (regressão logística, redes neurais e máquina de vetores de suporte) serão descritos nas Seções 5.1, 5.2 e 5.3.

5.1 Regressão Logística

Regressão logística é um modelo de classificador estatístico que mede a relação entre variáveis independentes e uma categoria (classe) (BISHOP, 2006).

5.1.1 Modelo

Segundo BISHOP (2006), considerando um problema de classificação de duas classes (\mathcal{C}_1 e \mathcal{C}_2), a probabilidade a posteriori da classe \mathcal{C}_1 pode ser escrita como a aplicação da função sigmoide

$$\sigma(z) = \frac{1}{1 + \exp(-z)} \quad (5.1)$$

sobre uma função linear do vetor de características $\phi = \phi(x)$ de tal forma que

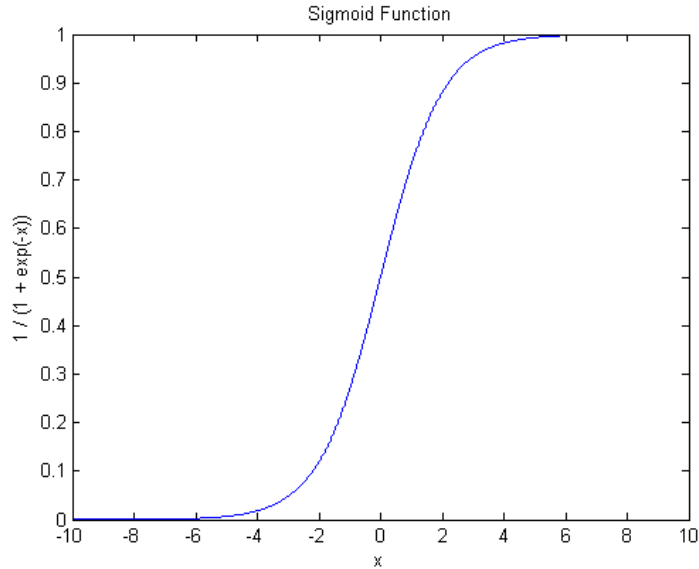
$$p(\mathcal{C}_1|\phi) = y(\phi) = \sigma(w^T \phi), \quad (5.2)$$

onde $p(\mathcal{C}_2|\phi) = 1 - p(\mathcal{C}_1|\phi)$, w^T é o vetor transposto de pesos (ou parâmetros) do modelo, x é o vetor de características de entrada para um determinado exemplo, e o vetor $\phi = \phi(x)$ é definido como

$$\phi_i(x) = \begin{cases} 1 & \text{se } i = 0 \\ x_i & \text{se } i \neq 0 \end{cases},$$

onde o termo $\phi_0(x) = 1$ é conhecido como unidade de viés. A Figura 5.1 mostra o formato da função sigmoide $\sigma(z)$.

Figura 5.1: Função sigmoide.



Para a determinação dos parâmetros do modelo de regressão logística, utiliza-se o método da máxima verossimilhança. Assim, para um conjunto de dados $\{\phi_n, t_n\}$ - onde

$$\phi_n = \phi_{ni}(x_n) = \begin{cases} 1 & \text{se } i = 0 \\ x_{ni} & \text{se } i \neq 0 \end{cases}$$

representa o vetor de características do exemplo n , e

$$t_n = \begin{cases} 1 & \text{se o exemplo } n \text{ pertence à classe } \mathcal{C}_1 \\ 0 & \text{se o exemplo } n \text{ pertence à classe } \mathcal{C}_2 \end{cases}$$

representa a classe do exemplo n -, com $n = 1, \dots, N$, a função de verossimilhança pode ser escrita como

$$p(t|w) = \prod_{n=1}^N y_n^{t_n} (1 - y_n)^{1-t_n}, \quad (5.3)$$

onde $t = (t_1, \dots, t_N)^T$, $y_n = p(\mathcal{C}_1|\phi_n)$, e N é o número de exemplos considerados. Pode-se, então, definir a função erro como o logaritmo negativo da verossimilhança

$$E(w) = -\ln(p(t|w)) = -\sum_{n=1}^N (t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n)), \quad (5.4)$$

onde $y_n = \sigma(w^T \phi_n)$ (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

O objetivo do método da máxima verossimilhança é encontrar valores para os parâmetros do modelo de forma a maximizar as probabilidades dos dados observados (exemplos), ou seja,

$$\max_{w \in \mathbb{R}^M} (p(t|w)) \quad (5.5)$$

ou, na prática,

$$\min_{w \in \mathbb{R}^M} (E(w)) = \min_{w \in \mathbb{R}^M} \left(- \sum_{n=1}^N \left(t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n) \right) \right), \quad (5.6)$$

onde M é o número de características consideradas.

Para a minimização da função erro, utiliza-se algoritmos de minimização, como, por exemplo, o método do gradiente descendente. Para este trabalho, por questões de desempenho, um algoritmo de minimização desenvolvido e registrado por Carl Edward Rasmussen e modificado por Andrew Y. Ng foi utilizado. O código fonte em MATLAB[®] deste algoritmo encontra-se no Anexo A.

Para a utilização destes algoritmos é, normalmente, necessário fornecer, além da função erro, o seu gradiente com respeito ao vetor de parâmetros (ou pesos) - uma vez que, para se minimizar uma função, deve-se encontrar o ponto onde suas derivadas são zero. Assim, ainda de acordo com BISHOP (2006) e HASTIE; TIBSHIRANI; FRIEDMAN (2009), o gradiente da função erro com respeito a w pode ser escrito como

$$\nabla E(w) = \sum_{n=1}^N (y_n - t_n) \phi_n. \quad (5.7)$$

5.1.2 Regularização

Regularização é uma técnica frequentemente utilizada para o controle do fenômeno de sobreajuste (ou *overfitting* em inglês). Esta técnica envolve a adição de uma penalidade à função erro de forma a evitar que coeficientes (ou parâmetros ou pesos) atinjam valores muito altos. Assim, a função erro a ser minimizada toma a forma

$$E(w) = E_D(w) + \lambda E_W(w), \quad (5.8)$$

onde λ é o coeficiente de regularização, que controla a importância relativa do erro dependente dos dados ($E_D(w)$) e do termo de regularização ($E_W(w)$) (BISHOP, 2006).

Segundo BISHOP (2006), uma das formas mais simples de regularização é a soma dos quadrados dos elementos do vetor de pesos (ou parâmetros). Neste caso,

$$E_W(w) = \frac{1}{2} w^T w, \quad (5.9)$$

onde $w^T w = w_1^2 + w_2^2 + \dots + w_M^2$, sendo M o número de características (ou parâmetros). É importante notar que o coeficiente w_0 - que serve de peso à característica constante unitária (unidade de viés) $\phi_{n0}(x_n) = 1$ - é omitido do termo regularizador (pode-se dizer que $w_0 = 0$ neste termo), pois a sua inclusão afeta a posição de origem dos dados (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Assim, para a regressão logística, substituindo (5.4) e (5.9) em (5.8), a função erro torna-se

$$E(w) = - \sum_{n=1}^N \left(t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n) \right) + \frac{\lambda}{2} w^T w \quad (5.10)$$

e o gradiente da função erro torna-se

$$\nabla E(w) = \sum_{n=1}^N (y_n - t_n) \phi_n + \lambda w. \quad (5.11)$$

A regularização permite que modelos complexos sejam treinados com conjuntos de dados de tamanho limitado sem que o problema de sobreajuste (*overfitting*) aconteça. No entanto, ao se utilizar regularização, o problema torna-se determinar um valor adequado para o coeficiente de regularização (λ) (BISHOP, 2006).

5.1.3 Predição

Como visto anteriormente, conforme BISHOP (2006), em um problema de classificação de duas classes (\mathcal{C}_1 e \mathcal{C}_2), a probabilidade a posteriori de uma classe \mathcal{C}_1 pode ser escrita da seguinte forma

$$p(\mathcal{C}_1|\phi) = y(\phi) = \sigma(w^T \phi), \quad (5.12)$$

onde $p(\mathcal{C}_2|\phi) = 1 - p(\mathcal{C}_1|\phi)$, $\sigma(\cdot)$ é a função sigmoide definida em (5.1), w^T é o vetor transposto de pesos (ou parâmetros) do modelo, x é o vetor de características de entrada para um determinado exemplo, e $\phi = \phi(x)$ é definido como

$$\phi_i(x) = \begin{cases} 1 & \text{se } i = 0 \\ x_i & \text{se } i \neq 0 \end{cases}.$$

Assim, para a predição da classe de um novo vetor de características ($\phi'(x')$), utiliza-se esta mesma equação. Portanto, a probabilidade de um novo vetor de características pertencer à classe \mathcal{C}_1 é dada por

$$p(\mathcal{C}_1|\phi') = y(\phi') = \sigma(w^T \phi'). \quad (5.13)$$

5.1.4 Classificação “Um Contra Todos”

Quando o problema de classificação exige mais de duas classes, utiliza-se um método conhecido como “um contra todos” (ou “*one versus all*” em inglês).

Este método de classificação baseia-se na utilização de K classificadores binários, onde K é o número de classes existentes. Cada um dos classificadores é criado de forma a ser capaz de diferenciar uma classe de todas as outras $K - 1$ classes (de onde vem o nome do método), ou seja, cada um dos classificadores é capaz de identificar se um determinado exemplo pertence à classe representada por si.

Assim, nesta arquitetura, para a realização de uma predição, utiliza-se o processo de predição comum (como visto na Seção 5.1.3) com cada um dos classificadores criados e, então, identifica-se qual destes classificadores apresenta a maior probabilidade a posteriori (o maior grau de certeza) sobre as características de entrada. Ou seja, prediz-se que um vetor de dados de entrada pertence a uma determinada classe se o classificador responsável por identificar exemplos desta classe for o que apresenta a maior probabilidade a posteriori em comparação com todos os outros classificadores.

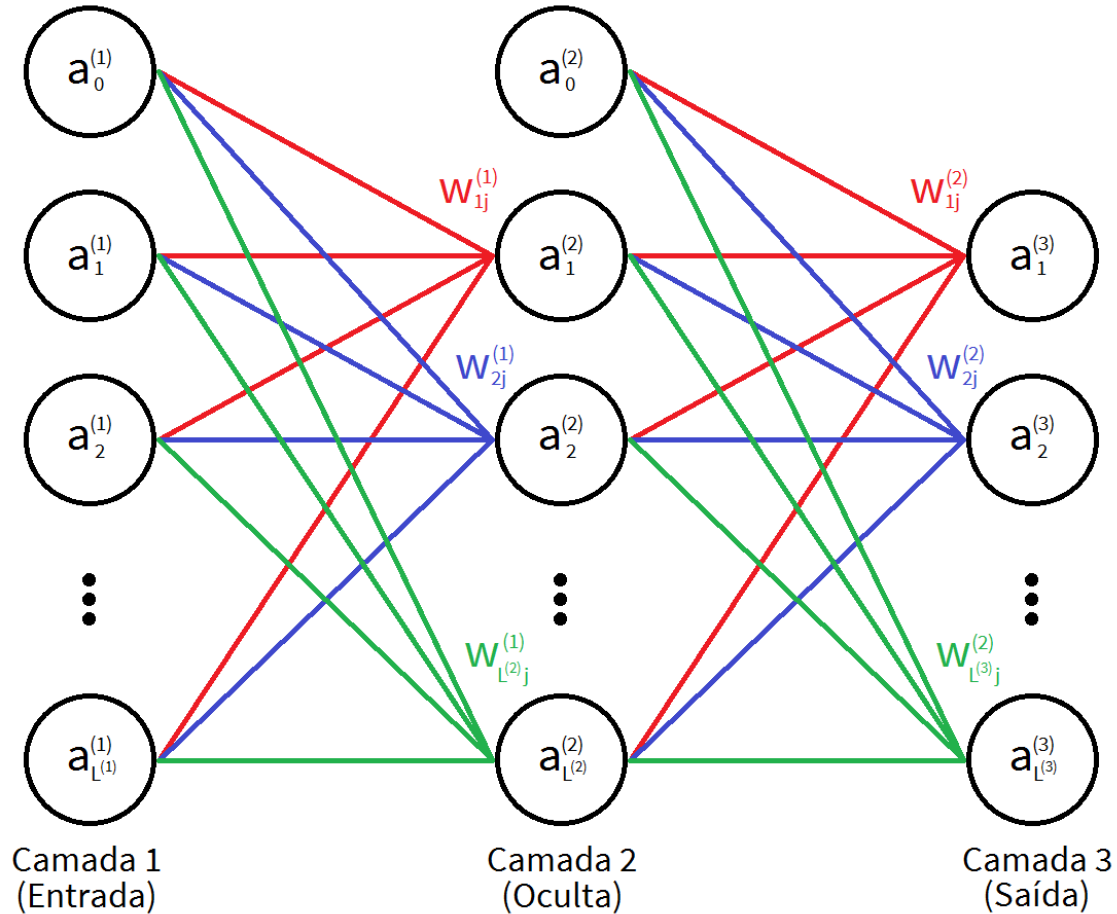
5.2 Redes Neurais

Redes Neurais são modelos de inteligência computacional capazes de aprender e reconhecer padrões. O mais bem sucedido modelo deste tipo no contexto de reconhecimento de padrões é a rede neural *feedforward*. As propriedades de aproximação das redes *feedforward* foram extensamente estudadas e descobertas como sendo muito gerais. Este tipo de rede pode ser aplicada tanto para classificação como para regressão. Redes neurais são, então, chamadas de aproximadores universais. Para muitas aplicações, o modelo pode ser significativamente mais compacto e, portanto, pode ser avaliado mais rapidamente

que uma máquina de vetores de suporte tendo o mesmo desempenho de generalização (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

A Figura 5.2 representa o modelo de rede neural utilizado neste projeto e pode ser utilizada como referência para o conteúdo desta seção.

Figura 5.2: Arquitetura da rede neural utilizada neste projeto.



5.2.1 Feedforward

Segundo BISHOP (2006), o modelo básico de rede neural pode ser descrito como uma série de transformações.

Primeiramente, define-se os valores de ativação da primeira camada (indicada pelo sobrescrito (1)) da rede - a camada de entrada - como

$$a_i^{(1)} = \begin{cases} 1 & \text{se } i = 0 \\ x_i & \text{se } i \neq 0 \end{cases},$$

onde x_i é a característica i do vetor de variáveis de entrada. O elemento $a_0^{(1)}$ é conhecido como unidade de viés da camada 1 (BISHOP, 2006).

Para a segunda camada (indicada pelo sobrescrito (2)) - a camada oculta -, define-se $z^{(2)}$, uma combinação linear das variáveis de entrada (valores de ativação da primeira

camada), como

$$z_i^{(2)} = \sum_{j=0}^{L^{(1)}} w_{ij}^{(1)} a_j^{(1)}, \quad (5.14)$$

onde $w_{ij}^{(1)}$ é uma matriz $L^{(2)} \times (L^{(1)} + 1)$ que representa os pesos de mapeamento da camada 1 para a camada 2, e $L^{(l)}$ é o número de unidades na camada l (excluindo a unidade de viés $a_0^{(l)}$). Os valores de ativação da camada 2 são uma transformação dos valores $z_i^{(2)}$

$$a_i^{(2)} = \begin{cases} 1 & \text{se } i = 0 \\ \sigma(z_i^{(2)}) & \text{se } i \neq 0 \end{cases},$$

onde $\sigma(\cdot)$ - a função de ativação da camada - é a função sigmoide, definida em (5.1). O elemento $a_0^{(2)}$ é a unidade de viés da camada 2 (BISHOP, 2006).

Finalmente, para a terceira camada (indicada pelo sobrescrito (3)) - a camada de saída -, define-se $z^{(3)}$, uma soma ponderada dos valores de ativação da camada 2, como

$$z_i^{(3)} = \sum_{j=0}^{L^{(2)}} w_{ij}^{(2)} a_j^{(2)}, \quad (5.15)$$

onde $w_{ij}^{(2)}$ é uma matriz $L^{(3)} \times (L^{(2)} + 1)$ que representa os pesos de mapeamento da camada 2 para a camada 3, e $L^{(l)}$ é o número de unidades na camada l (excluindo a unidade de viés $a_0^{(l)}$). Assim, os valores de ativação da camada 3 são

$$y_i = a_i^{(3)} = \sigma(z_i^{(3)}), \quad (5.16)$$

onde $\sigma(\cdot)$ - a função de ativação da camada - é a função sigmoide, definida em (5.1). É importante notar que a camada de saída possui $L^{(3)}$ unidades, onde $L^{(3)}$ é o número de classes sendo consideradas (BISHOP, 2006).

Este método é conhecido como *forward propagation* (BISHOP, 2006).

5.2.2 Treinamento

Considerando um caso de classificação binária com

$$t = \begin{cases} 1 & \text{se o exemplo pertence à classe } \mathcal{C}_1 \\ 0 & \text{se o exemplo pertence à classe } \mathcal{C}_2 \end{cases}$$

e com uma única unidade de saída - que tem como função de ativação a função sigmoide $y = a^{(3)} = \sigma(z^{(3)})$, definida em (5.1), de tal forma que $0 \leq y(x, w) \leq 1$. Pode-se interpretar $y(x, w)$ como a probabilidade condicional

$$y(x, w) = p(\mathcal{C}_1|x) = 1 - p(\mathcal{C}_2|x). \quad (5.17)$$

A distribuição condicional de um exemplo $\{x, t\}$ é, então, uma distribuição de Bernoulli da forma

$$p(t|x, w) = y(x, w)^t (1 - y(x, w))^{1-t}. \quad (5.18)$$

Para um conjunto de treinamento com exemplos independentes, a função erro, que é dada pelo logaritmo negativo da verossimilhança, é da forma

$$E(w) = - \sum_{n=1}^N \left(t_n \ln(y_n) + (1 - t_n) \ln(1 - y_n) \right), \quad (5.19)$$

onde $y_n = y(x_n, w)$ (BISHOP, 2006).

Para o caso de $K > 2$ classes, pode-se utilizar uma rede com $L^{(3)} = K$ unidades de saída (K classificadores binários) de função de ativação sigmoide. Associado a cada unidade de saída, existe um rótulo binário de classe

$$t_k = \begin{cases} 1 & \text{se o exemplo pertence à classe } \mathcal{C}_k \\ 0 & \text{se o exemplo não pertence à classe } \mathcal{C}_k \end{cases},$$

onde $k = 1, \dots, K$. Assumindo que, dado um vetor de entrada, os rótulos de classe são independentes, então a distribuição condicional é

$$p(t|x, w) = \prod_{k=1}^K y_k(x, w)^{t_k} (1 - y_k(x, w))^{1-t_k}. \quad (5.20)$$

A função erro, que é o logaritmo negativo da função de verossimilhança, compreendendo todos os exemplos de entrada, é

$$E(w) = - \sum_{n=1}^N \sum_{k=1}^K \left(t_{nk} \ln(y_{nk}) + (1 - t_{nk}) \ln(1 - y_{nk}) \right), \quad (5.21)$$

onde $y_{nk} = y_k(x_n, w)$ é o valor da unidade k da camada de saída para o exemplo n , t_{nk} é o valor esperado na unidade k da camada de saída para o exemplo n , N é o número de exemplos fornecidos como entrada para o treinamento da rede, e $K = L^{(3)}$ é o número de classes consideradas (número de unidades de saída) (BISHOP, 2006).

O treinamento da rede neural baseia-se em encontrar os parâmetros $w = \{w^{(1)}, w^{(2)}\}$ que minimizam a função erro, ou seja,

$$\min_w (E(w)). \quad (5.22)$$

Como a função erro é uma função contínua em w , o seu mínimo acontece no ponto que o gradiente da função é nulo, ou seja,

$$\nabla E(w) = 0. \quad (5.23)$$

Para a resolução desta equação, utiliza-se procedimentos numéricos iterativos, e muitos destes algoritmos utilizam o valor do gradiente da função erro. Como mostra a Seção 5.2.3, o método de retropropagação de erros (ou *error backpropagation*) é capaz de calcular $\nabla E(w)$ eficientemente (BISHOP, 2006).

Como no caso da regressão logística (Seção 5.1), para este trabalho, por questões de desempenho, um algoritmo de minimização desenvolvido e registrado por Carl Edward Rasmussen e modificado por Andrew Y. Ng foi utilizado. O código fonte em MATLAB[®] deste algoritmo encontra-se no Anexo A.

5.2.3 Retropropagação de Erros

A retropropagação (ou *backpropagation*) é um método eficiente para o cálculo do gradiente de uma função erro $E(w)$ de uma rede neural *feedforward*. Este método baseia-se em um esquema de troca de mensagens em que informação é enviada, alternadamente, para frente e para trás na rede (BISHOP, 2006).

O cálculo da função erro, como mostra (5.21), compreende uma soma de termos que pode ser escrita da seguinte forma

$$E(w) = \sum_{n=1}^N E_n(w), \quad (5.24)$$

onde N é o número de exemplos fornecidos como entrada para o treinamento da rede, e $E_n(w)$ é o valor da função erro para o exemplo n do conjunto de treinamento, que vale

$$E_n(w) = - \sum_{k=1}^K \left(t_{nk} \ln(y_{nk}) + (1 - t_{nk}) \ln(1 - y_{nk}) \right), \quad (5.25)$$

onde $y_{nk} = y_k(x_n, w)$ é o valor da unidade k da camada de saída para o exemplo n , t_{nk} é o valor esperado na unidade k da camada de saída para o exemplo n , e $K = L^{(3)}$ é o número de classes consideradas (número de unidades de saída) (BISHOP, 2006).

A derivada de $E_n(w)$ em respeito ao peso w_{ji} é, pela regra da cadeia,

$$\frac{\partial E_n(w)}{\partial w_{ji}^{(l-1)}} = \frac{\partial E_n(w)}{\partial z_{nj}^{(l)}} \frac{\partial z_{nj}^{(l)}}{\partial w_{ji}^{(l-1)}}, \quad (5.26)$$

onde $w_{ji}^{(l-1)}$ é o peso associado à conexão da unidade i da camada $l - 1$ com a unidade j da camada l , e $z_{nj}^{(l)}$ é o valor da unidade j da camada l para o exemplo n . Diz-se, então, por questão de simplificação, que

$$\delta_{nj}^{(l)} \equiv \frac{\partial E_n(w)}{\partial z_{nj}^{(l)}}. \quad (5.27)$$

Assim, sabendo que

$$\frac{\partial z_{nj}^{(l)}}{\partial w_{ji}^{(l-1)}} = a_{ni}^{(l-1)}, \quad (5.28)$$

tem-se, substituindo (5.27) e (5.28) em (5.26), que

$$\frac{\partial E_n(w)}{\partial w_{ji}^{(l-1)}} = \delta_{nj}^{(l)} a_{ni}^{(l-1)}. \quad (5.29)$$

Portanto, para a avaliação das derivadas, precisa-se calcular o valor de δ_{nj} para cada unidade oculta e unidade de saída da rede (BISHOP, 2006).

Conforme BISHOP (2006), para as unidades de saída, tem-se que

$$\delta_{nk}^{(3)} = y_{nk} - t_{nk}. \quad (5.30)$$

Já para as unidades ocultas, faz-se, pela regra da cadeia,

$$\delta_{nj}^{(2)} \equiv \frac{\partial E_n(w)}{\partial z_{nj}^{(2)}} = \sum_{k=1}^K \frac{\partial E_n(w)}{\partial z_{nk}^{(3)}} \frac{\partial z_{nk}^{(3)}}{\partial z_{nj}^{(2)}}. \quad (5.31)$$

Como, segundo (5.27),

$$\frac{\partial E_n(w)}{\partial z_{nk}^{(3)}} = \delta_{nk}^{(3)} \quad (5.32)$$

e

$$\frac{\partial z_{nk}^{(3)}}{\partial z_{nj}^{(2)}} = w_{kj}^{(2)} \sigma'(z_{nj}^{(2)}), \quad (5.33)$$

então, substituindo (5.32) e (5.33) em (5.31),

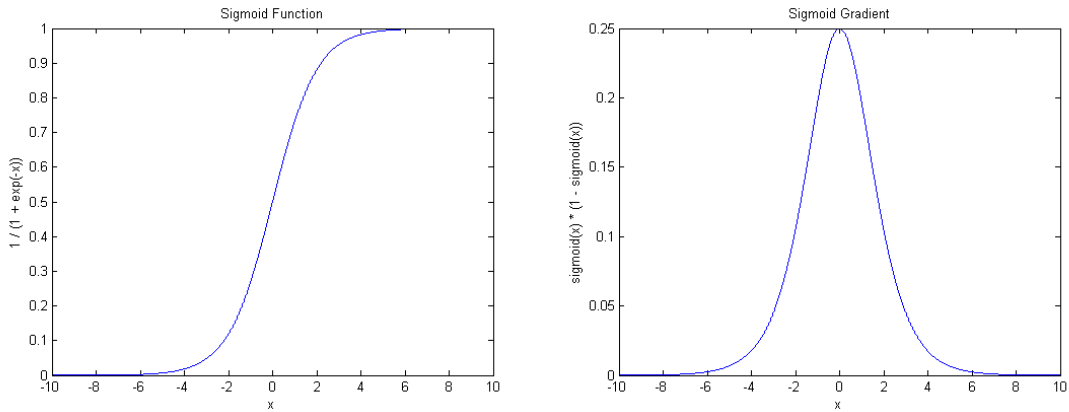
$$\delta_{nj}^{(2)} = \sigma'(z_{nj}^{(2)}) \sum_{k=1}^K w_{kj}^{(2)} \delta_{nk}^{(3)}, \quad (5.34)$$

onde

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \quad (5.35)$$

é o gradiente da função sigmoide. A Figura 5.3 mostra o formato da função sigmoide e de seu gradiente.

Figura 5.3: Função sigmoide (a) e seu gradiente (b).



(a) Função sigmoide.

(b) Gradiente da função sigmoide.

A derivada do erro total (considerando todos os N exemplos de entrada) é

$$\frac{\partial E(w)}{\partial w_{ji}^{(l)}} = \sum_{n=1}^N \frac{\partial E_n(w)}{\partial w_{ji}^{(l)}}. \quad (5.36)$$

5.2.4 Regularização

A regularização é uma forma de controlar a complexidade de uma rede neural de forma a evitar o sobreajuste (ou *overfitting* em inglês). Com a adição da regularização, a função erro toma a forma

$$E(w) = E_D(w) + \lambda E_W(w), \quad (5.37)$$

onde λ é o coeficiente de regularização, que controla a importância relativa do erro dependente dos dados ($E_D(w)$) e do termo de regularização ($E_W(w)$). Como já foi mencionado na Seção 5.1.2, o regularizador quadrático é o método mais simples de regularização e tem a forma

$$E_W(w) = \frac{1}{2} w^T w, \quad (5.38)$$

ou, para o caso da rede neural utilizada,

$$E_W(w) = \frac{1}{2} \left(\sum_{i=1}^{L^{(1)}} \sum_{j=1}^{L^{(2)}} \left(w_{ji}^{(1)} \right)^2 + \sum_{i=1}^{L^{(2)}} \sum_{j=1}^{L^{(3)}} \left(w_{ji}^{(2)} \right)^2 \right), \quad (5.39)$$

onde $L^{(l)}$ é o número de unidades na camada l (excluindo a unidade de viés). É importante notar que o coeficiente $w_{j0}^{(l)}$ - que serve de peso à característica constante unitária (unidade de viés) - é omitido do termo regularizador (pode-se dizer que $w_{j0}^{(l)} = 0$ neste termo), pois a sua inclusão afeta a posição de origem dos dados (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

O gradiente da função erro tem a forma

$$\nabla E(w) = \nabla E_D(w) + \lambda \nabla E_W(w), \quad (5.40)$$

onde $E_D(w)$ é obtido utilizando o método de retropropagação de erros (ou *error backpropagation*) - descrito na Seção 5.2.3 -, e o gradiente do termo regularizador é

$$\frac{\partial E_W(w)}{\partial w_{ji}^{(l)}} = w_{ji}^{(l)}. \quad (5.41)$$

5.2.5 Predição

Dado um vetor de entrada x , para realizar a predição da classe a qual este vetor pertence, aplica-se o *feedforward*, como visto na Seção 5.2.1, utilizando as matrizes peso $w^{(1)}$ e $w^{(2)}$ da rede neural treinada. O valor de cada uma das unidades de saída representa a probabilidade do vetor de entrada pertencer à classe representada por aquela unidade. Assim, a classe prevista é a classe representada pela unidade de saída com a maior probabilidade.

5.2.6 Valores Iniciais para os Pesos

A correta inicialização das matrizes de peso é uma etapa fundamental para o bom treinamento de uma rede neural. A inicialização com zeros (matrizes nulas) faz a rede tornar-se simétrica, e suas derivadas, nulas, o que impede o algoritmo de chegar a qualquer resultado. Já a inicialização dos pesos com valores muito grandes geralmente gera soluções ruins. Por isso, os valores iniciais para os pesos de uma rede neural são, geralmente, aleatórios e próximos de zero. Assim, o modelo inicia de forma quase linear (uma vez que a parte operacional da função sigmoide é praticamente linear com valores próximos a zero) e torna-se não-linear com o aumento dos pesos (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

5.2.7 Quantidade de Camadas e de Unidades em cada Camada

Uma rede neural possui três tipos de camadas: camada de entrada, camada oculta e camada de saída. Todo modelo contém, obrigatoriamente, uma camada de entrada e uma camada de saída. O número de camadas ocultas de uma rede neural é um valor arbitrário. A utilização de múltiplas camadas ocultas permite a construção de características mais complexas na rede. Geralmente, a quantidade de camadas ocultas utilizada é estabelecida através de experimentação (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

A primeira camada de uma rede neural é a camada de entrada, que corresponde ao vetor de entrada ($a_i^{(1)} = x_i$ para $i > 0$) mais uma unidade de viés ($a_0^{(1)} = 1$) e, portanto,

tem um número de unidades igual ao tamanho do vetor de entrada ($L^{(1)}$) mais um (unidade de viés). A última camada é a de saída, que corresponde às possíveis classes as quais uma entrada pode pertencer, portanto o número de unidades de saída é exatamente igual ao número de classes consideradas ($L^{(3)}$ na arquitetura deste projeto). As camadas entre a camada de entrada e a camada de saída são chamadas de ocultas. O número de unidades em uma camada oculta é arbitrário. Quanto menor o número de unidades ocultas, menor é a flexibilidade do modelo para capturar as não-linearidades dos dados. Por outro lado, se o número de unidades ocultas for muito grande, o modelo pode perder a capacidade de generalização para novos dados de entrada. Geralmente, prefere-se treinar uma rede neural com um grande número de unidades ocultas e regularização - que é capaz de controlar a complexidade do modelo, evitando a perda de capacidade de generalização (HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

5.3 Máquina de Vetores de Suporte

Máquinas de Vetores de Suporte são modelos de aprendizado supervisionado para a análise e reconhecimento de padrões. SVMs (*Support Vector Machines*) são utilizadas para classificação e regressão de dados e tentam resolver o problema através do conceito de margem, que é a menor distância entre a superfície de decisão e qualquer uma das amostras. Nestes modelos, a superfície de decisão (ou hiperplano) é escolhida de forma a ser aquela com a maior margem. Uma propriedade importante deste modelo de classificador é que a determinação dos parâmetros do modelo corresponde a um problema de otimização convexa, e, portanto, qualquer solução local é, também, um valor ótimo global (BISHOP, 2006).

5.3.1 Modelo

Por simplicidade, considera-se o modelo de classificação linear binário seguinte

$$y(x) = w^T \phi(x) + b, \quad (5.42)$$

onde $\phi(x)$ é uma transformação do vetor de características, w^T é o vetor transposto de pesos (ou parâmetros) do modelo, e $b = w_0$ é o parâmetro de viés, que se encontra explícito. Os dados de entrada são classificados de acordo com o sinal de $y(x)$. Considera-se, também, um conjunto de dados de treinamento da forma $\{x_n, t_n\}$ com $n = 1, \dots, N$, onde N é o número de exemplos de entrada, M é o número de características de um exemplo, $x_n \in \mathbb{R}^M$ representa o vetor de características do exemplo n , e $t_n \in \{1, -1\}$ representa a classe do exemplo n . Assume-se, momentaneamente, que o conjunto de dados de treinamento é linearmente separável no espaço de características, de forma que exista ao menos uma escolha de parâmetros (w e b) que satisfaça

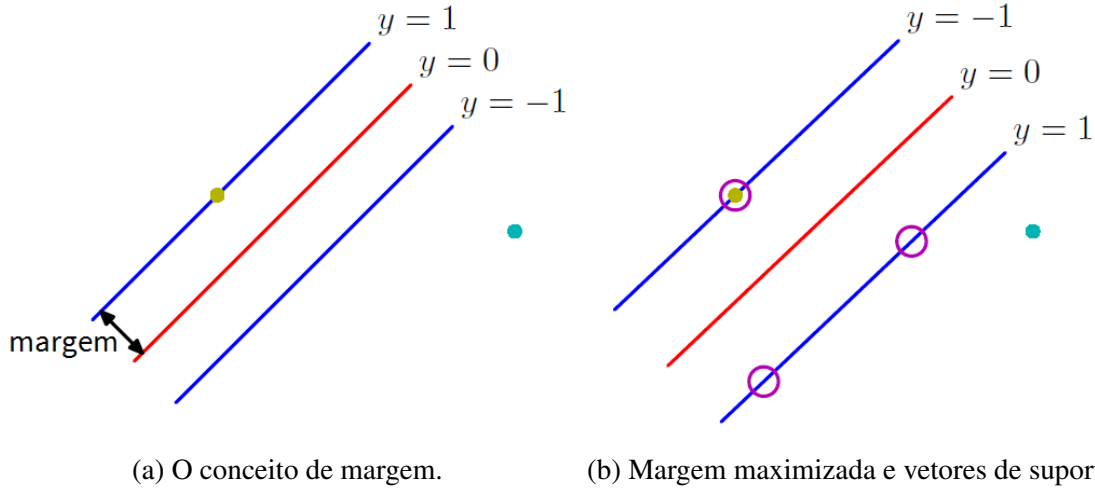
$$\begin{cases} y(x_n) > 0 & \text{se } t_n = +1 \\ y(x_n) < 0 & \text{se } t_n = -1 \end{cases}.$$

Desta forma,

$$t_n y(x_n) > 0 \quad (5.43)$$

para todos os vetores de treinamento (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Figura 5.4: A fronteira de decisão e o conceito de margem. Ao se maximizar a margem, obtém-se uma fronteira de decisão cuja localização é determinada por um subconjunto de pontos, conhecidos como vetores de suporte (indicados pelos círculos) (BISHOP, 2006).



Em máquinas de vetores de suporte, a fronteira de decisão é escolhida de forma a maximizar a margem - que é definida como a menor distância perpendicular entre a fronteira de decisão e qualquer um dos exemplos de entrada, como ilustrado na Figura 5.4.

Segundo BISHOP (2006), a distância perpendicular de um ponto x de um hiperplano definido por $y(x) = 0$ é dada por $|y(x)|/||w||$. Assim, para que todos os pontos estejam corretamente classificados - de forma que $t_n y(x_n) > 0$ para todo n -, a distância de um ponto x_n até a fronteira de decisão deve ser

$$\frac{t_n y(x_n)}{||w||} = \frac{t_n (w^T \phi(x_n) + b)}{||w||}. \quad (5.44)$$

Assim, segundo BISHOP (2006) e HASTIE; TIBSHIRANI; FRIEDMAN (2009), para se maximizar o tamanho da margem, deve-se otimizar os parâmetros w e b da seguinte forma

$$\arg \max_{w,b} \left(\frac{1}{||w||} \min_n \left(t_n (w^T \phi(x_n) + b) \right) \right), \quad (5.45)$$

onde se maximiza o tamanho da margem para o ponto mais próximo da fronteira de decisão. A solução direta para este problema é muito complexa, portanto transforma-se este problema em uma versão equivalente de resolução mais simples. Como a distância entre um ponto x_n e a fronteira de decisão não muda com a multiplicação de w e de b por uma mesma constante K

$$\frac{t_n (K w^T \phi(x_n) + K b)}{||K w||} = \frac{t_n (w^T \phi(x_n) + b)}{||w||}, \quad (5.46)$$

pode-se dizer que

$$t_n (w^T \phi(x_n) + b) = 1 \quad (5.47)$$

para o ponto mais próximo à fronteira (fixa-se o valor mínimo do numerador). Desta forma, todos os pontos devem obedecer

$$t_n (w^T \phi(x_n) + b) \geq 1, n = 1, \dots, N. \quad (5.48)$$

Por definição, sempre haverá um ponto mais próximo, portanto sempre existirá um valor de n para o qual $t_n(w^T \phi(x_n) + b) = 1$. Uma vez a margem maximizada, existirão ao menos dois valores de n para os quais $t_n(w^T \phi(x_n) + b) = 1$. Assim, o problema de otimização (5.45) passa a ser

$$\arg \max_{w,b} \frac{1}{\|w\|}, \quad (5.49)$$

o que é equivalente a

$$\arg \min_{w,b} \frac{1}{2} \|w\|^2, \quad (5.50)$$

onde o fator $1/2$ é incluído por conveniência (e não altera o resultado). Compactamente, o problema de otimização é

$$\begin{cases} \arg \min_{w,b} \frac{1}{2} \|w\|^2 \\ t_n(w^T \phi(x_n) + b) \geq 1, \quad n = 1, \dots, N \end{cases} .$$

Resolve-se tal problema de otimização utilizando o método dos multiplicadores de Lagrange. Assim, introduz-se os multiplicadores $a_n \geq 0$ com $n = 1, \dots, N$, de tal forma que a função lagrangiana seja

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{n=1}^N a_n (t_n(w^T \phi(x_n) + b) - 1), \quad (5.51)$$

onde $a = (a_1, \dots, a_N)^T$. Com $\nabla L(w, b, a) = 0$, obtém-se

$$w = \sum_{n=1}^N a_n t_n \phi(x_n) \quad (5.52)$$

e

$$0 = \sum_{n=1}^N a_n t_n. \quad (5.53)$$

Portanto, utilizando (5.52) e (5.53) em (5.51), tem-se a representação dual do problema da máxima margem, no qual se maximiza

$$\tilde{L}(a) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N a_n a_m t_n t_m k(x_n, x_m) \quad (5.54)$$

com respeito a a sujeito às restrições

$$a_n \geq 0, \quad n = 1, \dots, N \quad (5.55)$$

e

$$\sum_{n=1}^N a_n t_n = 0, \quad (5.56)$$

onde, nesta situação simplificada, $k(x, x') = \phi(x)^T \phi(x')$ é o função *kernel* (ou núcleo) do modelo. Este é um problema mais simples, que pode ser resolvido por *softwares* padrões. É importante ressaltar que outras funções *kernel* podem ser utilizadas, e estas serão vistas na Seção 5.3.2 (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Pode-se mostrar que um problema de otimização desta forma satisfaz as condições de Karush-Kuhn-Tucker (KKT), que, neste caso, são

$$a_n \geq 0, \quad (5.57)$$

$$t_n y(x_n) - 1 \geq 0 \quad (5.58)$$

e

$$a_n (t_n y(x_n) - 1) = 0. \quad (5.59)$$

Logo, para cada exemplo x_n , ou $a_n = 0$ ou $t_n y(x_n) = 1$. Os pontos em que $a_n \neq 0$ são chamados de vetores de suporte e estão situados sobre os hiperplanos de margem máxima, como é mostrado na Figura 5.4. Uma vez que o modelo esteja treinado, somente os vetores de suporte precisam ser mantidos, pois são os únicos que afetam a predição de novos vetores de entrada (já que $a_n = 0$ para os outros pontos). Ainda, para se obter o valor de b , resolve-se (5.59) para qualquer vetor de suporte (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Até agora, assumiu-se que o conjunto de dados de entrada é linearmente separável. Na prática, as distribuições de pontos de cada uma das classes pode-se sobrepor. Neste caso, o modelo proposto até aqui não seria capaz de ter uma boa generalização para vetores de entrada não vistos no conjunto de dados de treinamento (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Portanto, conforme BISHOP (2006) e HASTIE; TIBSHIRANI; FRIEDMAN (2009), precisa-se modificar o modelo proposto de forma que ele aceite que alguns pontos do conjunto de treinamento sejam mal classificados, pois assim se consegue criar um modelo com um melhor poder de generalização. Para isto, modifica-se o sistema de forma a permitir que existam pontos no “lado errado” das fronteiras, mas apresentando uma penalização que aumenta com a distância da fronteira. É conveniente, para o problema de otimização, que esta penalidade seja uma função linear da distância. Assim, adiciona-se ao sistema variáveis de folga (*slack variables* em inglês) $\xi_n \geq 0$, onde $n = 1, \dots, N$. Estas variáveis são definidas como

$$\xi_n = \begin{cases} 0 & \text{se o ponto está sobre ou dentro dos limites corretos} \\ |t_n - y(x_n)| & \text{caso o contrário} \end{cases}$$

Assim, um ponto que está sobre a fronteira de decisão $y(x_n) = 0$ terá $\xi_n = 1$, e pontos com $\xi_n > 1$ estarão classificados de forma incorreta. A restrição (5.48) do modelo anterior torna-se, então,

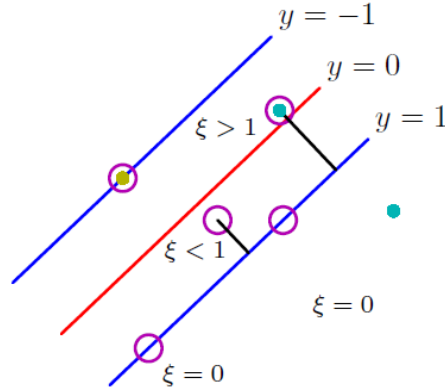
$$t_n y(x_n) \geq 1 - \xi_n, \quad n = 1, \dots, N, \quad (5.60)$$

onde as variáveis de folga devem respeitar $\xi_n \geq 0$. Assim, pontos para os quais $\xi_n = 0$ estão corretamente classificados e estão ou sobre a margem ou no lado correto da margem; pontos para os quais $0 < \xi_n \leq 1$ estão dentro da margem, mas do lado correto da fronteira de decisão; e pontos para os quais $\xi_n > 1$ estão do lado errado da fronteira de decisão e estão classificados incorretamente, como mostra a Figura 5.5.

O objetivo do modelo é a maximização da margem, penalizando os pontos que se encontram do lado errado da margem. Portanto, modificando (5.50) do modelo anterior, minimiza-se

$$\arg \min_{w,b} C \sum_{n=1}^N \xi_n + \frac{1}{2} \|w\|^2, \quad (5.61)$$

Figura 5.5: Variáveis de folga $\xi_n \geq 0$. Pontos com círculos são vetores de suporte (BISHOP, 2006).



onde $C > 0$ regula a importância da penalidade dada a uma variável de folga e da margem, ou seja, da minimização dos erros de treinamento e da complexidade do modelo. O parâmetro C é análogo ao inverso do coeficiente de regularização dos outros modelos ($C \equiv 1/\lambda$) (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Assim, conforme BISHOP (2006) e HASTIE; TIBSHIRANI; FRIEDMAN (2009), o problema de minimização torna-se

$$\begin{cases} \arg \min_{w,b} C \sum_{n=1}^N \xi_n + \frac{1}{2} \|w\|^2 \\ t_n y(x_n) \geq 1 - \xi_n, \quad n = 1, \dots, N \\ \xi_n \geq 0 \end{cases} .$$

E, introduzindo os multiplicadores de Lagrange $a_n \geq 0$ e $\mu_n \geq 0$, a função lagrangiana é

$$L(w, b, a) = \frac{1}{2} \|w\|^2 + C \sum_{n=1}^N \xi_n - \sum_{n=1}^N a_n (t_n y(x_n) - 1 + \xi_n) - \sum_{n=1}^N \mu_n \xi_n. \quad (5.62)$$

O conjunto correspondente de condições KKT é

$$a_n \geq 0, \quad (5.63)$$

$$t_n y(x_n) - 1 + \xi_n \geq 0, \quad (5.64)$$

$$a_n (t_n y(x_n) - 1 + \xi_n) = 0, \quad (5.65)$$

$$\mu_n \geq 0, \quad (5.66)$$

$$\xi_n \geq 0 \quad (5.67)$$

e

$$\mu_n \xi_n = 0, \quad (5.68)$$

onde $n = 1, \dots, N$.

Assim, otimizando-se w , b e $\{\xi_n\}$ e fazendo uso da definição de $y(x)$ (5.42), obtém-se

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w = \sum_{n=1}^N a_n t_n \phi(x_n), \quad (5.69)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow 0 = \sum_{n=1}^N a_n t_n \quad (5.70)$$

e

$$\frac{\partial L}{\partial \xi_n} = 0 \Rightarrow a_n = C - \mu_n. \quad (5.71)$$

E, finalmente, utilizando estes resultados em (5.62), obtém-se a representação dual da função lagrangiana na forma

$$\tilde{L}(a) = \sum_{n=1}^N a_n - \frac{1}{2} \sum_{n=1}^N \sum_{m=1}^N a_n a_m t_n t_m k(x_n, x_m), \quad (5.72)$$

que é idêntica à representação dual do modelo anterior, porém as restrições tornam-se

$$\sum_{n=1}^N a_n t_n = 0. \quad (5.73)$$

e, como $a_n \geq 0$, $\mu_n \geq 0$ e $a_n = C - \mu_n$, então

$$0 \leq a_n \leq C, \quad n = 1, \dots, N. \quad (5.74)$$

Para a resolução do modelo, maximiza-se a representação dual da função lagrangiana.

Novamente, após a maximização, os pontos para os quais $a_n \neq 0$ são os vetores de suporte. Estes pontos devem, portanto, satisfazer (5.65), ou

$$t_n y(x_n) = 1 - \xi_n. \quad (5.75)$$

Se $a_n < C$, então $\mu_n > 0$ e $\xi_n = 0$, e, portanto, estes pontos estão sobre a margem. Pontos para os quais $a_n = C$ estão dentro da margem e podem estar corretamente classificados se $\xi_n \leq 1$ ou incorretamente classificados se $\xi_n > 1$.

Para a maximização da representação dual da função lagrangiana (obtenção dos valores de a_n e de b), utiliza-se algoritmos de otimização especializados. Para este trabalho, por questões de desempenho, utilizou-se uma versão simplificada do algoritmo *Sequential Minimal Optimization* (SMO).

5.3.2 Função Kernel

Segundo HASTIE; TIBSHIRANI; FRIEDMAN (2009), os tipos de função *kernel* de máquinas de vetores de suporte mais populares na literatura são:

Polinômio de Grau d : $k(x, x') = (1 + \langle x, x' \rangle)^d$

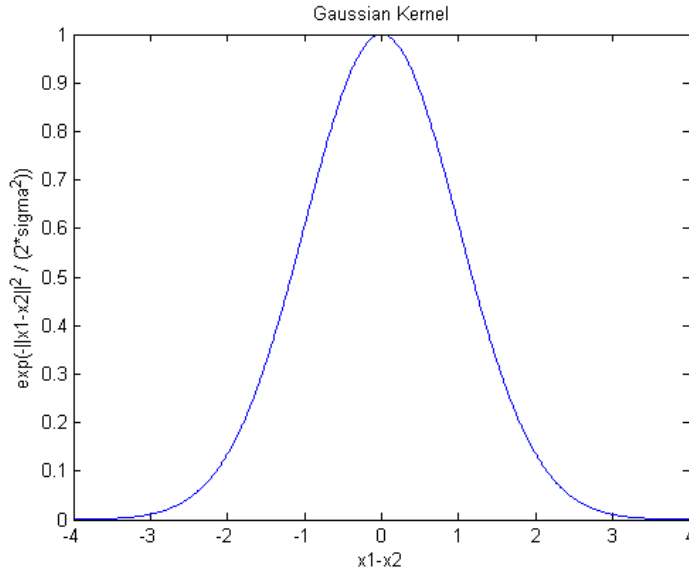
Função de Base Radial (RBF): $k(x, x') = \exp(-\gamma \|x - x'\|^2) = \exp(-\|x - x'\|^2 / (2\sigma^2))$

Rede Neural: $k(x, x') = \tanh(k_1 \langle x, x' \rangle + k_2)$

onde d é o grau do polinômio do núcleo polinomial, γ e σ são constantes arbitrárias do núcleo de função de base radial, e k_1 e k_2 são constantes arbitrárias do núcleo de rede neural.

Neste trabalho, a função *kernel* utilizada no modelo de máquinas de vetores de suporte é o núcleo gaussiano, ou função de base radial. Esta função serve como uma medida de similaridade entre dois pontos. A forma desta função pode ser vista na Figura 5.6.

Figura 5.6: Núcleo gaussiano (ou função de base radial) com $\sigma = 1$.



5.3.3 Predição

Segundo BISHOP (2006) e HASTIE; TIBSHIRANI; FRIEDMAN (2009), para a classificação de novos vetores de entrada, avalia-se o sinal de $y(x)$ - definido em (5.42) -, o que, expresso em termos dos parâmetros a_n e da função *kernel* pela substituição de w por (5.69), é

$$y(x) = \sum_{n=1}^N a_n t_n k(x, x_n) + b. \quad (5.76)$$

Observa-se, aqui, novamente, que não é mais necessário o conhecimento dos pontos que não são vetores de suporte para a predição de um novo vetor de entrada - uma vez que para um exemplo n que não é um vetor de suporte, $a_n = 0$. Assim, diz-se que a previsão é

$$\text{sign}(y(x)) = \begin{cases} +1 & \text{se o exemplo pertence à classe } \mathcal{C}_1 \\ -1 & \text{se o exemplo pertence à classe } \mathcal{C}_2 \end{cases}.$$

5.3.4 Classificação “Um Contra Todos”

Existem diversos métodos para a extensão do modelo de máquina de vetores de suporte a problemas de $K > 2$ classes. Na prática, o mais utilizado é o método conhecido como “um contra todos” (ou *one versus all* em inglês) - que, apesar de apresentar algumas limitações, está entre os que necessitam de menor capacidade computacional (BISHOP, 2006; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Este método baseia-se na construção de K classificadores, um para cada classe. O classificador $k = 1, \dots, K$ é treinado usando os exemplos pertencentes à classe \mathcal{C}_k como exemplos positivos e os exemplos pertencentes as outras $K - 1$ classes como exemplos negativos. Assim, o classificador k é capaz de identificar se um determinado exemplo pertence à classe \mathcal{C}_k (BISHOP, 2006).

Um dos problemas da utilização deste método é que, em máquinas de vetores de suporte, a saída do modelo não representa uma probabilidade (as únicas saídas possíveis são $+1$ ou -1). Por isso, existe a possibilidade de mais de um dos K classificadores

responderem positivamente para uma determinada entrada (ocasião na qual não é possível responder, objetivamente, a qual das classes a entrada fornecida pertence) (BISHOP, 2006).

Assim, nesta arquitetura, para a realização de uma predição, utiliza-se o processo de predição comum (como visto na Seção 5.3.3) com cada um dos classificadores criados e, então, identifica-se qual destes classificadores apresenta uma resposta positiva sobre as características de entrada. Ou seja, prediz-se que um vetor de dados de entrada pertence a uma determinada classe se o classificador responsável por identificar exemplos desta classe for o único a responder positivamente para este vetor de entrada. Caso mais de um classificador responda positivamente para um vetor de entrada, prediz-se uma das classes representadas por estes classificadores.

5.4 Validação Cruzada

Uma forma de estimar o melhor valor para os parâmetros de cada um dos modelos (como os parâmetros de regularização e o número de unidades ocultas de uma rede neural) é através de validação cruzada (ou *cross-validation* em inglês). Este procedimento consiste em utilizar parte do conjunto de dados disponível para a avaliação da performance do modelo para vetores de entrada nunca vistos em treinamento quando utilizando diferentes valores para os parâmetros. Desta forma, pode-se escolher os valores de parâmetros que resultam na melhor performance para os conjuntos de dados de treinamento e de validação cruzada. Assim, obtém-se um modelo complexo o suficiente para ajustar-se ao conjunto de treinamento, porém flexível o suficiente para conseguir generalizar para um conjunto de dados nunca visto até o momento (o conjunto de dados utilizado no procedimento de validação cruzada) (BISHOP, 2006).

Os parâmetros escolhidos durante o procedimento de *cross-validation* podem, no entanto, causar um sobreajuste (*overfitting*) no conjunto de validação cruzada. Por esta razão, utiliza-se um terceiro conjunto de dados - o conjunto de teste - para a avaliação da real performance de um modelo para dados de entrada nunca vistos (nem em treinamento, nem em validação cruzada) (BISHOP, 2006).

É importante notar que os dados utilizados na validação cruzada não pertencem nem ao conjunto de treinamento - que é utilizado para o treinamento do modelo - nem ao conjunto de teste - que é utilizado para a avaliação da performance do modelo para dados de entrada nunca vistos (BISHOP, 2006).

5.5 Métricas de Qualidade de Classificação

É importante, após o desenvolvimento de um modelo de aprendizado, que se consiga avaliar, numericamente, a qualidade do algoritmo. Métricas de qualidade de classificação tentam avaliar a capacidade de tomada de decisão de um algoritmo. As seguintes métricas são utilizadas para a avaliação dos modelos: taxa de acerto, valor preditivo positivo, taxa de verdadeiras positivas, taxa de verdadeiras negativas, valor preditivo negativo, *F1 score*, *informedness* e *markedness* (SCHRÖDER; THIELE; LEHNER, 2011).

5.5.1 Matriz de Confusão

A matriz de confusão é um método para a contabilização de predições de um modelo. Ela é montada de forma a classificar as predições de um algoritmo em quatro tipos:

Verdadeiro Positivo: representa as predições positivas que eram realmente positivas;

Falso Positivo: representa as predições positivas que não eram positivas na realidade;

Falso Negativo: representa as predições negativas que não eram negativas na realidade;

Verdadeiro Negativo: representa as predições negativas que eram realmente negativas.

A Tabela 5.1 ilustra a matriz de confusão para o caso binário (NILSON, 2014).

Tabela 5.1: Matriz de confusão para o caso binário.

		Realidade	
		Positivo	Negativo
Predição	Positivo	Verdadeiro Positivo (TP)	Falso Positivo (FP)
	Negativo	Falso Negativo (FN)	Verdadeiro Negativo (TN)

Em particular, no caso deste trabalho, para o classificador k da classe \mathcal{C}_k , um verdadeiro positivo é um caso corretamente classificado como pertencente à classe \mathcal{C}_k ; um verdadeiro negativo é um caso corretamente classificado como não pertencente à classe \mathcal{C}_k ; um falso positivo é um caso incorretamente classificado como pertencente à classe \mathcal{C}_k ; e um falso negativo é um caso incorretamente classificado como não pertencente à classe \mathcal{C}_k .

5.5.2 Taxa de Acerto

A taxa de acerto é a proporção de casos preditos corretamente. O cálculo da taxa de acerto é realizado conforme a Equação (5.77):

$$TA = \frac{TP + TN}{N}, \quad (5.77)$$

onde TP indica a quantidade de verdadeiros positivos, TN indica a quantidade de verdadeiros negativos, e N indica o número total de casos classificados (FAWCETT, 2006; POWERS, 2011).

5.5.3 Valor Preditivo Positivo

O valor preditivo positivo (ou *precision*, como é chamado em inglês) é a proporção de casos preditos como positivos que são realmente positivos, ou seja, este valor estima a probabilidade de um caso predito como positivo ser realmente positivo. O cálculo do valor preditivo positivo é realizado conforme a Equação (5.78):

$$VPP = \frac{TP}{TP + FP}, \quad (5.78)$$

onde TP indica a quantidade de verdadeiros positivos, e FP indica a quantidade de falsos positivos (FAWCETT, 2006; POWERS, 2011; SCHRÖDER; THIELE; LEHNER, 2011).

5.5.4 Taxa de Verdadeiros Positivos

Segundo FAWCETT (2006), POWERS (2011) e SCHRÖDER; THIELE; LEHNER (2011), a taxa de verdadeiros positivos (ou *recall*, como é chamado em inglês) é a proporção de casos positivos reais que foram corretamente classificados, ou seja, este valor

estima a probabilidade de um caso realmente positivo ser previsto como positivo. O cálculo da taxa de verdadeiros positivos é realizado conforme a Equação (5.79):

$$TVP = \frac{TP}{TP + FN}, \quad (5.79)$$

onde TP indica a quantidade de verdadeiros positivos, e FN indica a quantidade de falsos negativos.

5.5.5 Taxa de Verdadeiros Negativos

A taxa de verdadeiros negativos (ou *inverse recall*, como é chamado em inglês) é a proporção de casos negativos reais que foram corretamente classificados, ou seja, este valor estima a probabilidade de um caso realmente negativo ser previsto como negativo. O cálculo da taxa de verdadeiros negativos é realizado conforme a Equação (5.80):

$$TVN = \frac{TN}{TN + FP}, \quad (5.80)$$

onde TN indica a quantidade de verdadeiros negativos, e FP indica a quantidade de falsos positivos (FAWCETT, 2006; POWERS, 2011; SCHRÖDER; THIELE; LEHNER, 2011).

5.5.6 Valor Preditivo Negativo

Segundo FAWCETT (2006), POWERS (2011) e SCHRÖDER; THIELE; LEHNER (2011), o valor preditivo negativo (ou *inverse precision*, como é chamado em inglês) é a proporção de casos preditos como negativos que são realmente negativos, ou seja, este valor estima a probabilidade de um caso predito como negativo ser realmente negativo. O cálculo do valor preditivo negativo é realizado conforme a Equação (5.81):

$$VPN = \frac{TN}{TN + FN}, \quad (5.81)$$

onde TN indica a quantidade de verdadeiros negativos, e FN indica a quantidade de falsos negativos.

5.5.7 F1 Score

Segundo FAWCETT (2006), POWERS (2011) e SCHRÖDER; THIELE; LEHNER (2011), o *F1 score* (ou *F1-measure*) é considerado, na área de inteligência artificial, como uma medida única da eficácia de um sistema. Esta medida combina o valor preditivo positivo (Seção 5.5.3) e a taxa de verdadeiros positivos (Seção 5.5.4) em uma única métrica. No entanto, esta medida ignora, completamente, os casos corretamente classificados como negativos (verdadeiros negativos). O cálculo do *F1 score* é realizado conforme a Equação (5.82):

$$F1 = \frac{2TP}{2TP + FP + FN} = \frac{2}{\frac{1}{VPP} + \frac{1}{TVP}}, \quad (5.82)$$

onde TP indica a quantidade de verdadeiros positivos, FP indica a quantidade de falsos positivos, FN indica a quantidade de falsos negativos, VPP é o valor preditivo positivo (como especificado na Seção 5.5.3), e TVP é a taxa de verdadeiros positivos (como especificado na Seção 5.5.4).

5.5.8 *Informedness*

Conforme POWERS (2011) e SCHRÖDER; THIELE; LEHNER (2011), *informedness* quantifica o quão informadas são as classificações de um preditor (em comparação a predições de “sorte”). Esta medida combina a taxa de verdadeiros positivos (Seção 5.5.4) com a taxa de verdadeiros negativos (Seção 5.5.5) em uma única métrica. O cálculo da *informedness* é realizado conforme a Equação (5.83):

$$Informedness = TVP + TVN - 1, \quad (5.83)$$

onde *TVN* é a taxa de verdadeiros negativos (como especificado na Seção 5.5.5), e *TVP* é a taxa de verdadeiros positivos (como especificado na Seção 5.5.4).

5.5.9 *Markedness*

Conforme POWERS (2011) e SCHRÖDER; THIELE; LEHNER (2011), *markedness* quantifica o quão óbvias as predições são (em comparação a predições de “sorte”). Esta medida combina o valor preditivo positivo (Seção 5.5.3) com o valor preditivo negativo (Seção 5.5.6) em uma única métrica. O cálculo do *markedness* é realizado conforme a Equação (5.84):

$$Markedness = VPP + VPB - 1, \quad (5.84)$$

onde *VPB* é o valor preditivo negativo (como especificado na Seção 5.5.6), e *VPP* é o valor preditivo positivo (como especificado na Seção 5.5.3).

6 RESUMO DE TRABALHOS ANTERIORES

Diversos estudos são realizados anualmente sobre a caracterização de movimentos de partes do corpo utilizando eletromiografia. Além do segmento do corpo considerado, estes trabalhos se diferenciam, principalmente, pelas metodologias e técnicas utilizadas na captura, interpretação ou classificação dos sinais.

GLETTE *et al.* (2008) utilizaram 4 pares de eletrodos (de resolução de 14 bits e a uma taxa de amostragem de 6 kHz) – posicionados especialmente para cada sujeito, nos quatro lados do antebraço, de forma a obter o melhor sinal possível – com a referência no pulso e várias técnicas de classificação diferentes (*k-Nearest-Neighbor*, *Decision Trees*, *Support Vector Machines*, *Embedded Cartesian Genetic Programming Evolvable Hardware* e *Functional Unit Row Evolvable Hardware Architecture*) para a identificação de até 8 movimentos diferentes (abertura, fechamento, flexão, extensão, desvio ulnar, desvio radial, pronação e supinação), obtendo diferentes taxas de acerto para as diferentes técnicas utilizadas. A Tabela 6.1 mostra a relação entre o método de classificação utilizado e a taxa de erros do classificador. O experimento é realizado em 3 dias, e os resultados incluem os resultados individuais de cada dia (Dia 1, Dia 2 e Dia 3), os resultados da combinação dos 3 dias (Dias 1-3) e o resultado da utilização dos dados de dois dias para o treinamento e do terceiro dia para o teste (2 de 3).

Tabela 6.1: Relação entre o método de classificação utilizado e a taxa de erros do classificador (GLETTE *et al.*, 2008).

	Dia 1	Dia 2	Dia 3	Dias 1-3	2 de 3
<i>K-Nearest-Neighbor</i>	3,5 %	4,6 %	4,6 %	4,5 %	5,6 %
<i>Decision Trees</i>	9,7 %	11,3 %	10,5 %	9,0 %	15,9 %
<i>Support Vector Machines</i>	4,2 %	4,0 %	2,6 %	4,5 %	5,4 %
<i>Embedded Cartesian Genetic Programming Evolvable Hardware</i>	9,8 %	4,0 %	5,3 %	9,0 %	10,6 %
<i>Functional Unit Row Evolvable Hardware Architecture</i>	9,0 %	4,6 %	4,0 %	4,9 %	8,4 %

BOSCHMANN *et al.* (2009) utilizaram 4 pares de eletrodos (de resolução de 24 bits e a uma taxa de amostragem de 2048 Hz) – posicionados especialmente para cada sujeito, nos quatro lados do antebraço, de forma a obter o melhor sinal possível – com a referência no pulso e máquina de vetores de suporte para a identificação de até 11 movimentos diferentes (extensão, flexão, desvio ulnar, desvio radial, pronação, supinação, abertura, fechamento, pegada de chave, pegada em pinça e extração do dedo indicador), obtendo diferentes taxas de acerto para diferentes quantidades de movimentos. A Ta-

bela 6.2 mostra a relação entre a quantidade de movimentos considerados e a taxa de acerto do classificador (os movimentos descartados são aqueles com as piores taxas de acerto).

Tabela 6.2: Relação entre a quantidade de movimentos considerados e a taxa de acerto do classificador (BOSCHMANN *et al.*, 2009).

Quantidade de Movimentos	11	10	9	8	7	6	5	4-2
Taxa de Acerto [%]	91,3	94,5	96,1	97,5	98,1	98,6	99,7	100,0

TANG *et al.* (2012) utilizaram 6 pares de eletrodos (de resolução de 10 bits e a uma taxa de amostragem de 1000 Hz) - arranjados na forma de uma rede de tamanho ajustável ao braço e posicionados na parte posterior do antebraço (capturando sinais dos músculos *extensor digitorum*, *extensor pollicis longus*, *extensor pollicis brevis*, *extensor indicis* e *extensor digiti minimi*) - e um conjunto de 6 classificadores LDA (análise discriminante linear) em cascata para a identificação de 11 poses distintas. A taxa de acerto média obtida no teste de 6 sujeitos foi de 94,8 %.

ATZORI *et al.* (2012) utilizaram 10 pares de eletrodos ativos duplamente diferenciais (a uma taxa de amostragem de 100 Hz) – 8 eletrodos posicionados, de forma igualmente espaçada, logo abaixo do cotovelo, a uma distância fixa da articulação radioumeral, e 2 eletrodos posicionados nos músculos flexor e extensor – e máquina de vetores de suporte para a identificação de até 52 movimentos diferentes (12 movimentos básicos dos dedos, 8 configurações isométricas e isotônicas da mão, 9 movimentos básicos do pulso e 23 movimentos funcionais e de agarramento), obtendo diferentes taxas de acerto para diferentes quantidades de movimentos. A Tabela 6.3 mostra a relação entre a quantidade de movimentos considerados e a taxa de erros do classificador.

Tabela 6.3: Relação entre a quantidade de movimentos considerados e a taxa de erros do classificador (ATZORI *et al.*, 2012).

Quantidade de Movimentos	52	3
Taxa de Erros [%]	20,3 ± 6,9	7,5 ± 7,2

GIJSBERTS *et al.* (2014) utilizaram a mesma base de dados empregada neste trabalho - a base de dados NinaPro (versão 2), que foi descrita no Capítulo 3 - e *Kernel Regularized Least Squares* (KRLS) - um algoritmo similar à máquina de vetores de suporte - com três núcleos diferentes (linear, RBF e $\exp-\chi^2$) para a classificação de 40 poses (imagens 1 a 40 da Figura 3.2), obtendo diferentes taxas de acerto para configurações distintas. A Tabela 6.4 mostra a relação entre o núcleo (*kernel*) de classificação utilizado, o conjunto de características utilizado e a taxa de acerto do classificador.

Tabela 6.4: Relação entre o núcleo (*kernel*) de classificação utilizado, o conjunto de características utilizado e a taxa de acerto do classificador (GIJSBERTS *et al.*, 2014).

	Linear	RBF	$\exp-\chi^2$
<i>RMS</i>	49,35 %	72,2 %	73,62 %
<i>sEMG Histogram</i>	63,1 %	74,48 %	75,94 %
<i>Marginal Discrete Wavelet Transform</i>	56,66 %	75,41 %	77 %
<i>RMS, Histogram, Wavelet Transform</i>	-	-	77,48 %

7 METODOLOGIA EXPERIMENTAL

7.1 Plataforma de Software

A plataforma de software utilizada para o desenvolvimento deste projeto é o MATLAB[®] da empresa MathWorks[®].

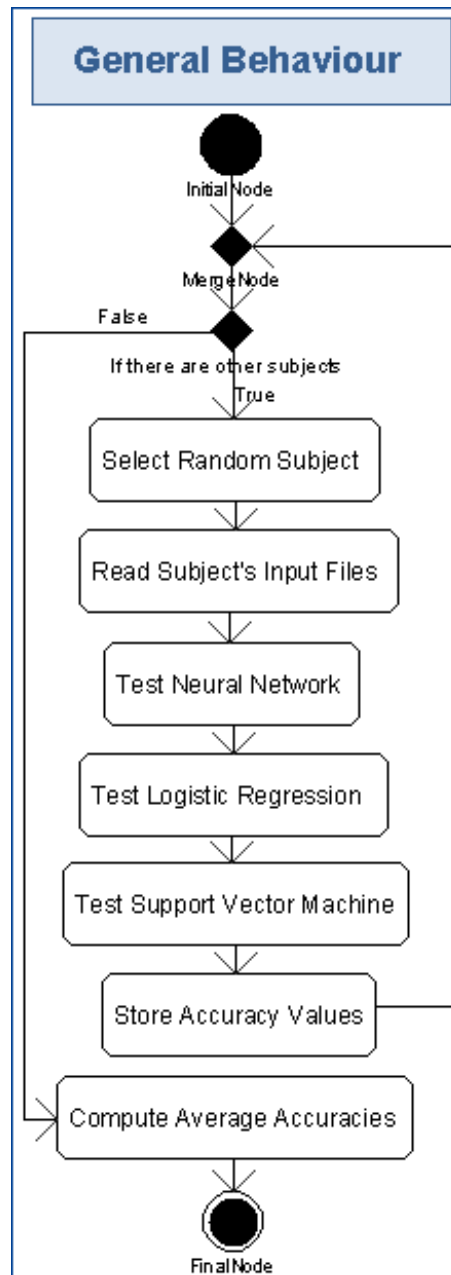
O MATLAB[®] (ou *Matrix Laboratory*) é, além de um ambiente de computação numérica multiparadigma, uma linguagem de programação interpretada de alto nível. A plataforma é conhecida por oferecer manipulação fácil e eficiente de matrizes e criação de gráficos científicos, no entanto a ferramenta conta com diversas outras funcionalidades, como o processamento de sinais e de imagens. As suas principais aplicações são: solução de sistemas complexos de equações, modelagem, simulação, prototipação e análise, exploração e visualização de dados.

Na prática, no caso deste projeto, esta plataforma de software age como uma facilitadora dos processos de desenvolvimento e teste de protótipos de sistemas numéricos, resolvendo estes tipos de problemas em uma fração do tempo que seria utilizado para a resolução do mesmo problema em linguagens de mais baixo nível, como C, C++, Fortran ou Java.

7.2 Funcionamento do Sistema

De uma maneira geral, o sistema segue o seguinte algoritmo. Primeiramente, um sujeito é selecionado de forma aleatória (porém sem possibilidade de repetição) da lista de sujeitos. Em seguida, a partição da base de dados que se refere ao sujeito selecionado é lida (uma descrição detalhada deste passo encontra-se na Seção 7.2.1). Ao finalizar a leitura dos dados do sujeito, uma lista de exemplos de treinamento - que podem ser utilizados como dados de entrada para algoritmos de aprendizado - já está disponível. Então, o sistema realiza os testes de desempenho nos modelos de aprendizado (uma descrição detalhada destes testes encontra-se na Seção 7.2.2) - uma rede neural, um modelo de regressão logística e uma máquina de vetores de suporte, nesta ordem. Quando todos os três testes foram realizados, o sistema armazena os resultados (dados de desempenho dos modelos) para futuro tratamento. Esta sequência de passos é repetida até que um número previamente especificado de sujeitos tenha sido processado ou até que todos os sujeitos do banco de dados tenham sido processados. Ao término do processamento de novos sujeitos, o sistema utiliza os dados de desempenho de cada um deles para calcular as medidas de desempenho médias de cada um dos algoritmos de aprendizado. A Figura 7.1 mostra um diagrama de fluxo do comportamento geral do sistema.

Figura 7.1: Diagrama de fluxo do comportamento geral do sistema.

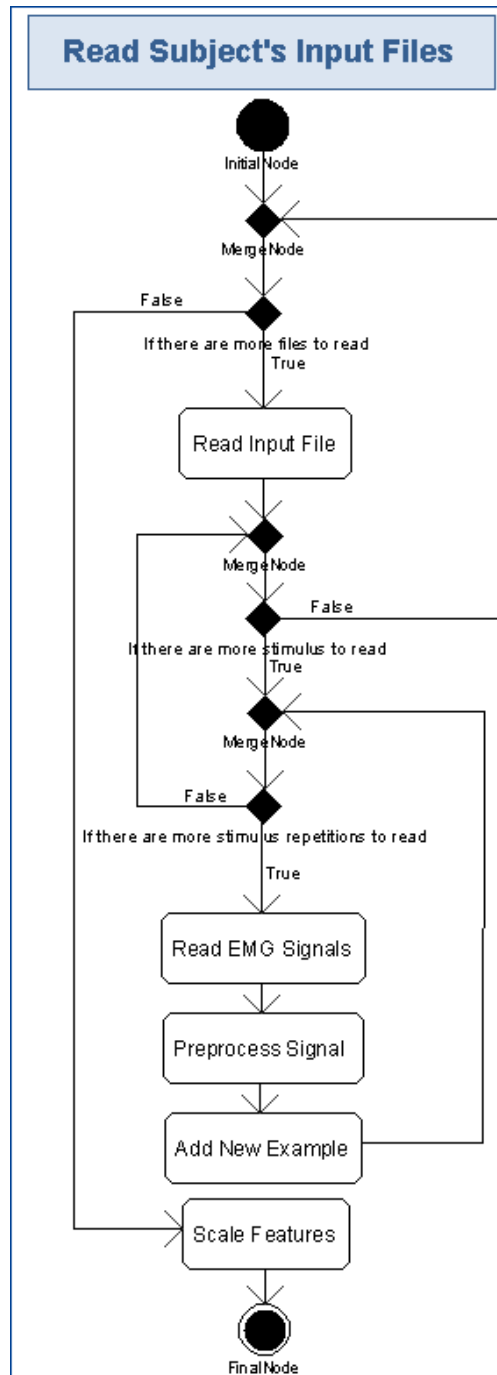


7.2.1 Leitura dos Dados de Entrada de um Sujeito

Na base de dados NinaPro, cada sujeito possui três arquivos de dados (um para cada exercício). Por isso, na etapa de leitura de dados de entrada, o algoritmo seguinte é repetido até três vezes (dependendo de quais exercícios estão sendo considerados). Inicialmente, o sistema realiza a leitura do arquivo correspondente a um dos exercícios considerados, o que traz ao ambiente de execução todas as variáveis descritas na Seção 3.2. Em seguida, para cada estímulo (movimento) considerado e para cada uma das repetições desse movimento (a base de dados contém 6 repetições de cada movimento), o sistema identifica os sinais EMG a serem processados, realiza o pré-processamento deles (esta etapa é melhor descrita na Seção 7.2.1.1) e, finalmente, possuindo as características das ondas (que são descritas na Seção 4.5), adiciona um novo exemplo (que poderá ser, futuramente, utilizado como dado de entrada para os modelos de aprendizado) na lista de

exemplos. Após finalizar a leitura e o processamento dos arquivos que contém os exercícios, o sistema dispõe de uma lista completa de exemplos de movimentos do sujeito em questão. O último passo é, então, o *Feature Scaling* (que é descrito na Seção 4.6). A Figura 7.2 mostra um diagrama de fluxo da leitura de dados de entrada de um sujeito.

Figura 7.2: Diagrama de fluxo da leitura dos dados de um sujeito.

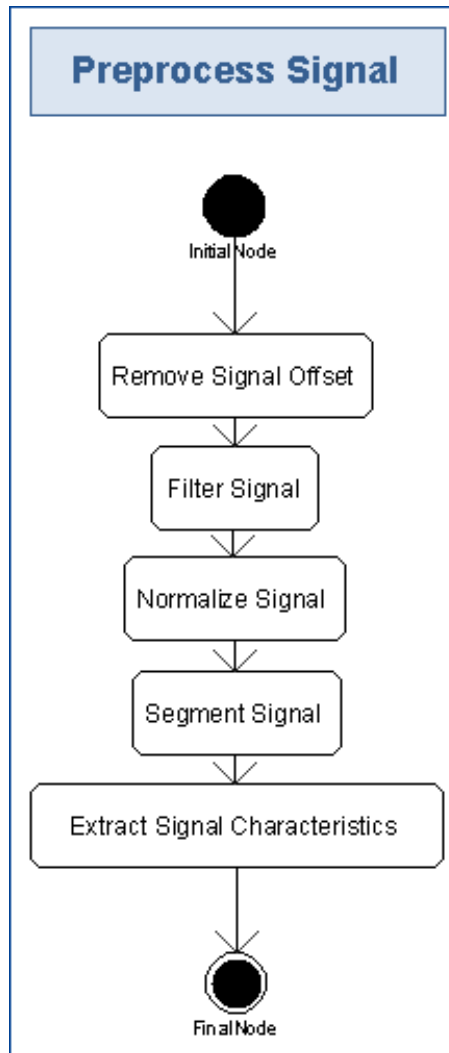


7.2.1.1 Pré-processamento do Sinal

A rotina de pré-processamento do sinal segue os seguintes passos: ajuste de *offset*, filtragem de sinal, normalização de sinal, segmentação de sinal e extração de características. O Capítulo 4 contém uma descrição detalhada de cada um desses passos. A Figura 7.3

mostra um diagrama de fluxo da etapa de pré-processamento do sinal. Como etapa de *feature scaling* só pode ser executada após a obtenção de todos os exemplos, ela não aparece neste diagrama e sim no diagrama de leitura de dados de entrada (veja a Figura 7.2).

Figura 7.3: Diagrama de fluxo do pré-processamento de um sinal.



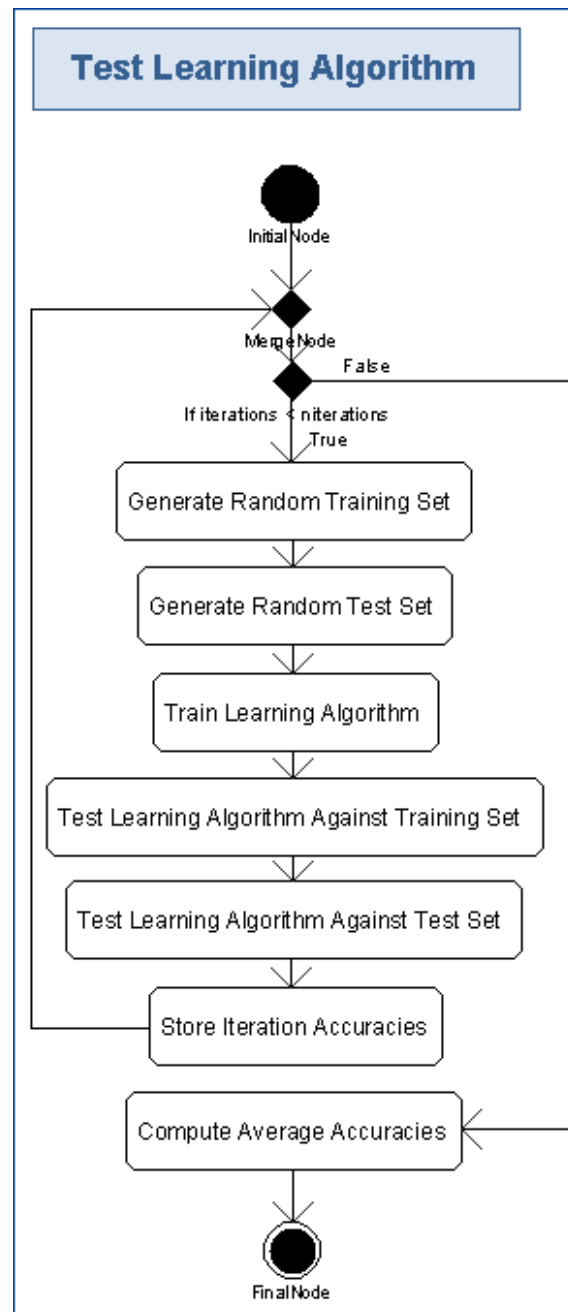
7.2.2 Teste de Algoritmo de Aprendizado

Como a geração do conjunto de exemplos de treinamento (*training set*) e de teste (*test set*) é aleatória (a Seção 7.2.3 contém uma explicação sobre cada um dos conjuntos), o teste de cada um dos algoritmos de aprendizado é repetido 100 vezes (número escolhido arbitrariamente). Assim, ao fim dos testes, pode-se obter taxas de acerto médias.

A rotina de teste inicia com a geração aleatória dos conjuntos de treinamento e de teste (a Seção 7.2.3.1 contém uma descrição detalhada do procedimento de geração de conjuntos). Em seguida, o sistema treina o algoritmo de aprendizado - que pode ser uma rede neural, um modelo de regressão logística ou uma máquina de vetores de suporte - com o conjunto de treinamento (*training set*). Assim que o treinamento é finalizado, o modelo é testado contra o conjunto de treinamento e contra o conjunto de teste, obtendo-se, ao fim dos testes, as medidas de desempenho para cada um dos conjuntos. Finalmente, as medidas de desempenho são armazenadas para futuro tratamento. Ao fim da execução dos

testes, o sistema calcula as medidas de desempenho médias para cada um dos conjuntos (treinamento e teste). A Figura 7.4 mostra um diagrama de fluxo da etapa de teste de um algoritmo de aprendizado.

Figura 7.4: Diagrama de fluxo do teste de um algoritmo de aprendizado.

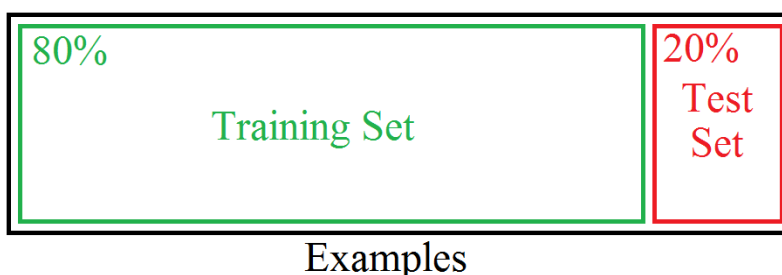


7.2.3 Conjuntos de Treinamento e de Teste

Durante a fase de teste, o conjunto de exemplos (criado na etapa de leitura de dados de entrada de um sujeito, descrito na Seção 7.2.1) é dividido em um conjunto de treinamento (*training set*) e um conjunto de teste (*test set*). O conjunto de treinamento é um subconjunto do conjunto de exemplos que será utilizado como entrada para o treinamento de um algoritmo de aprendizado. Além disso, após a fase de treinamento, este conjunto é utilizado para verificar o desempenho do modelo contra o conjunto de treinamento, ou

seja, a capacidade do modelo de classificar, corretamente, exemplos já conhecidos. Em sua configuração padrão, este conjunto contém 80 % das entradas do conjunto de exemplos. O conjunto de teste é, por sua vez, um subconjunto do conjunto de exemplos que é utilizado para verificar o desempenho do modelo quando classificando exemplos desconhecidos (não utilizados na fase de treinamento). Em sua configuração padrão, este conjunto contém 20 % das entradas do conjunto de exemplos. A Figura 7.5 mostra a divisão do conjunto de exemplos.

Figura 7.5: Divisão do conjunto de exemplos entre os conjuntos de treinamento (verde) e de teste (vermelho).



7.2.3.1 Geração dos Conjuntos

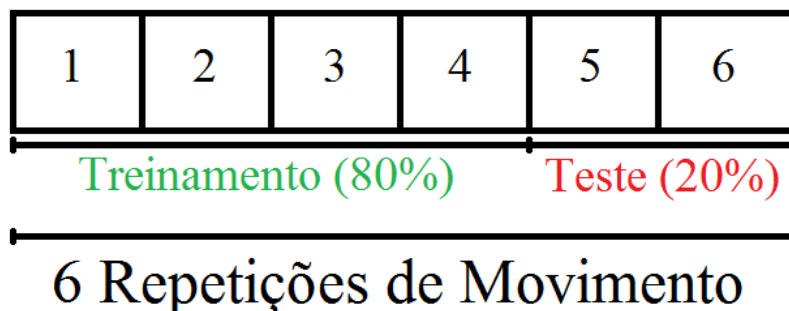
A geração dos conjuntos de treinamento e de teste a partir do conjunto de exemplos é feita de forma aleatória. No entanto, o algoritmo de partição garante que 80 % (configuração padrão) dos exemplos de cada um dos movimentos (repetições) estarão no conjunto de treinamento e que os outros 20 % (configuração padrão) estarão no conjunto de teste.

Como visto anteriormente, cada movimento considerado pelo projeto NinaPro possui 6 repetições (veja a Figura 7.6) - ou seja, no momento da aquisição de dados, os sujeitos repetem cada um dos movimentos 6 vezes. Assim, o algoritmo de partição do conjunto de exemplos utiliza a seguinte estratégia para a aleatorização dos exemplos. Para cada movimento, o sistema seleciona os dados referentes às 6 repetições daquele movimento e as embaralha. Em seguida, as primeiras 4 repetições (80 % arredondado para baixo) são adicionadas ao conjunto de exemplos de treinamento, e as últimas 2 (20 % arredondado para cima) são adicionadas ao conjunto de exemplos de teste. A Figura 7.7 mostra a divisão das 6 repetições entre os conjuntos de treinamento e de teste.

Figura 7.6: Seis repetições em cada movimento.

1	2	3	4	5	6	Movimento 1
1	2	3	4	5	6	Movimento 2
1	2	3	4	5	6	...
1	2	3	4	5	6	Movimento N

Figura 7.7: Divisão dos exemplos de um mesmo movimento entre os conjuntos de treinamento e de teste.



7.3 Determinação de Parâmetros dos Modelos

Como visto na Seção 5.4, os valores ideais para os parâmetros dos modelos de aprendizado utilizados podem ser estimados através do método de validação cruzada (ou *cross-validation* em inglês). Nesta técnica, parte dos dados de treinamento é utilizada como exemplo de teste para a avaliação do desempenho do modelo para diferentes valores de parâmetros. O conjunto de parâmetros que obtiver a melhor taxa de acertos sobre os conjuntos de treinamento e de validação cruzada é selecionado como o novo conjunto de parâmetros do modelo.

No caso deste projeto, como o número de exemplos para cada um dos movimentos é muito pequeno (6 repetições por movimento) e aproveitando do fato que a quantidade de sujeitos contidos na base de dados NinaPro é elevada (40 sujeitos), um desses quarenta sujeitos foi utilizado como “conjunto de validação cruzada”.

O procedimento de teste ocorre como descrito na Seção 7.2, porém apenas um sujeito é utilizado. Os modelos são treinados com 80 % dos exemplos, e o conjunto de teste (que agora é interpretado como um conjunto de validação cruzada) é composto dos outros 20 % dos exemplos. Uma série de parâmetros é testada, e aqueles que obtêm as melhores taxas de acerto são selecionados. Mais detalhadamente, estes parâmetros são escolhidos usando uma busca em grade, onde $\lambda \in \{0,001; 0,003; 0,01; 0,03; 0,1; 0,3; 1\}$, $L^{(2)} \in \{10; 20; 30; \dots; 100\}$, $\sigma \in \{0,001; 0,003; 0,01; 0,03; 0,1; 0,3; 1; 1,3; 2\}$, e $C \in \{0,01; 0,1; 1; \dots; 10\,000\}$.

No modelo de regressão logística, existe apenas um valor a ser estimado: o parâmetro de regularização (λ). Este parâmetro foi estimado com o valor ideal de $\lambda = 0,1$ ao atingir uma taxa de acerto de 100 %, tanto para o conjunto de treinamento quanto para o conjunto de validação cruzada. O modelo de rede neural possui dois valores a serem estimados: o parâmetro de regularização (λ) e o número de unidades ocultas ($L^{(2)}$). Da mesma forma que no caso do modelo de regressão logística, esses parâmetros foram determinados ao alcançarem uma taxa de acerto de 100 % para ambos os conjuntos. Obteve-se, então $\lambda = 0,01$ e $L^{(2)} = 50$. O modelo de máquina de vetores de suporte, por sua vez, possui outros dois parâmetros a serem estimados: o parâmetro de regularização ($C \equiv 1/\lambda$) e o parâmetro do núcleo gaussiano sendo utilizado (σ). Novamente, esses parâmetros foram estimados ao alcançarem uma taxa de acerto de 100 % para ambos os conjuntos como sendo $C = 10$ e $\sigma = 1$.

8 RESULTADOS E DISCUSSÕES

Este capítulo tem o intuito de apresentar o desempenho dos três modelos de aprendizado considerados (regressão logística, redes neurais e máquina de vetores de suporte) para uma série de configurações distintas. O primeiro conjunto de configurações visa avaliar o efeito da utilização de conjuntos variados de eletrodos no desempenho dos modelos (veja a Seção 8.1). O segundo conjunto de configurações tem o objetivo de identificar o desempenho dos algoritmos para conjuntos de características diversos (veja a Seção 8.2). O terceiro e último conjunto de configurações tem o intuito de evidenciar a qualidade dos modelos para tamanhos de conjunto de treinamento diferentes (veja a Seção 8.3). É importante ressaltar que os valores de desempenho exibidos neste capítulo são os valores médios dos desempenhos de 39 dos 40 sujeitos contidos na base de dados NinaPro (um dos sujeitos não é considerado nesta etapa, uma vez que ele foi utilizado durante a validação cruzada para estimar os parâmetros dos modelos, conforme explicado na Seção 7.3) e foram obtidos seguindo o procedimento descrito na Seção 7.2.

8.1 Desempenho Variando o Conjunto de Eletrodos

Nesta seção, avalia-se o desempenho dos modelos de aprendizado de máquina para conjuntos variáveis de eletrodos. Os seguintes conjuntos de eletrodos são avaliadas:

- conjunto de 8 eletrodos igualmente espaçados em volta do antebraço mais 4 eletrodos em músculos selecionados (Seção 8.1.1);
- conjunto de 4 eletrodos igualmente espaçados em volta do antebraço mais 4 eletrodos em músculos selecionados (Seção 8.1.2);
- conjunto de 4 eletrodos em músculos selecionados (Seção 8.1.3).
- conjunto de 8 eletrodos igualmente espaçados em volta do antebraço (Seção 8.1.4);
- conjunto de 4 eletrodos igualmente espaçados em volta do antebraço (Seção 8.1.5);
- conjunto de 2 eletrodos igualmente espaçados em volta do antebraço (Seção 8.1.6);

Após a exibição dos resultados, na Seção 8.1.7, faz-se uma avaliação de desempenho geral, comparando todos os conjuntos considerados.

8.1.1 Utilizando 8 Eletrodos em Volta do Antebraço e 4 Eletrodos em Músculos Selecionados

Nesta seção, considera-se os sinais de todos os eletrodos disponíveis na base de dados NinaPro (eletrodos 1 a 12, conforme mostra a Figura 3.1).

8.1.1.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,3%. A Tabela 8.1 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.1: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,3	97,8	95,1	98,6	99,3	98,7
Taxa de Verdadeiras Positivas [%]	94,8	93,9	95,0	97,5	99,1	97,3
Taxa de Verdadeiras Negativas [%]	98,3	99,4	98,6	99,6	99,8	99,6
Valor Preditivo Negativo [%]	99,0	98,8	99,0	99,5	99,8	99,5
<i>F1 Score</i> [%]	93,4	94,7	93,8	97,5	99,0	97,5
<i>Informedness</i> [%]	93,1	93,3	93,6	97,1	98,9	97,0
<i>Markedness</i> [%]	93,4	96,7	94,2	98,2	99,1	98,3

8.1.1.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,3%. A Tabela 8.2 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.2: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,8	97,6	94,8	98,7	99,1	98,9
Taxa de Verdadeiras Positivas [%]	94,6	93,6	95,0	97,4	99,2	97,9
Taxa de Verdadeiras Negativas [%]	98,5	99,4	98,5	99,6	99,7	99,7
Valor Preditivo Negativo [%]	99,0	98,8	99,1	99,5	99,8	99,6
<i>F1 Score</i> [%]	93,5	94,6	93,7	97,5	98,9	97,9
<i>Informedness</i> [%]	93,1	93,1	93,5	97,0	98,9	97,6
<i>Markedness</i> [%]	93,8	96,4	93,8	98,2	98,9	98,5

8.1.1.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 92,2%. A Tabela 8.3 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.3: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	86,8	92,5	92,5	96,7	98,1	98,1
Taxa de Verdadeiras Positivas [%]	94,9	88,9	89,5	92,4	94,9	92,9
Taxa de Verdadeiras Negativas [%]	95,5	98,3	98,1	99,2	99,6	99,7
Valor Preditivo Negativo [%]	99,0	97,9	98,1	98,6	99,0	98,7
<i>F1 Score</i> [%]	88,8	88,9	88,8	93,1	95,5	94,1
<i>Informedness</i> [%]	90,4	87,2	87,7	91,7	94,4	92,6
<i>Markedness</i> [%]	85,9	90,5	90,6	95,3	97,2	96,9

8.1.2 Utilizando 4 Eletrodos em Volta do Antebraço e 4 Eletrodos em Músculos Selecionados

Nesta seção, considera-se os sinais de 8 dos 12 eletrodos disponíveis na base de dados NinaPro. Em particular, considera-se apenas 4 dos 8 eletrodos igualmente espaçados em volta do antebraço (eletrodos 1, 3, 5 e 7, conforme mostra a Figura 3.1) e todos os 4 eletrodos de posição anatômico precisa (eletrodos 9 a 12, conforme mostra a Figura 3.1).

8.1.2.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 93,7%. A Tabela 8.4 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.4: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,3	94,1	92,4	97,0	98,5	97,1
Taxa de Verdadeiras Positivas [%]	92,4	90,7	92,4	94,1	97,6	95,0
Taxa de Verdadeiras Negativas [%]	97,7	98,5	98,0	99,2	99,5	99,2
Valor Preditivo Negativo [%]	98,5	98,3	98,6	98,9	99,5	99,1
<i>F1 Score</i> [%]	90,8	90,8	91,2	94,4	97,5	95,1
<i>Informedness</i> [%]	90,1	89,2	90,5	93,4	97,2	94,3
<i>Markedness</i> [%]	90,9	92,4	91,0	96,0	98,1	96,2

8.1.2.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 94,0%. A Tabela 8.5 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.5: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,2	94,5	92,2	97,1	98,2	97,3
Taxa de Verdadeiras Positivas [%]	92,2	91,6	91,8	94,2	98,4	95,5
Taxa de Verdadeiras Negativas [%]	98,0	98,6	98,0	99,2	99,4	99,2
Valor Preditivo Negativo [%]	98,5	98,4	98,5	98,9	99,7	99,2
<i>F1 Score</i> [%]	91,1	91,7	90,6	94,5	97,9	95,5
<i>Informedness</i> [%]	90,3	90,3	89,8	93,5	97,8	94,8
<i>Markedness</i> [%]	91,7	92,9	90,7	96,0	97,9	96,5

8.1.2.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 88,7%. A Tabela 8.6 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.6: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço e 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	83,9	88,1	89,4	94,1	96,5	96,2
Taxa de Verdadeiras Positivas [%]	92,1	86,4	84,1	88,4	92,1	89,1
Taxa de Verdadeiras Negativas [%]	94,6	96,8	97,7	98,6	99,1	99,4
Valor Preditivo Negativo [%]	98,4	97,5	97,1	97,9	98,5	98,0
<i>F1 Score</i> [%]	85,6	84,8	84,3	89,2	92,8	90,4
<i>Informedness</i> [%]	86,7	83,3	81,8	87,1	91,3	88,5
<i>Markedness</i> [%]	82,4	85,6	86,5	92,0	95,1	94,3

8.1.3 Utilizando 4 Eletrodos em Músculos Selecionados

Nesta seção, considera-se os sinais de 4 dos 12 eletrodos disponíveis na base de dados NinaPro. Em particular, considera-se apenas os 4 eletrodos de posição anatômico precisa (eletrodos 9 a 12, conforme mostra a Figura 3.1).

8.1.3.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 84,0%. A Tabela 8.7 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.7: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	83,1	84,0	82,9	87,1	88,8	94,8
Taxa de Verdadeiras Positivas [%]	81,2	81,0	81,9	81,9	87,7	90,5
Taxa de Verdadeiras Negativas [%]	95,6	96,3	95,8	97,0	97,2	98,7
Valor Preditivo Negativo [%]	96,5	96,4	96,6	96,6	97,7	98,2
<i>F1 Score</i> [%]	80,0	80,7	80,4	82,1	86,5	90,8
<i>Informedness</i> [%]	76,9	77,3	77,8	79,0	84,9	89,2
<i>Markedness</i> [%]	79,6	80,5	79,6	83,8	86,6	93,1

8.1.3.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 84,6%. A Tabela 8.8 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.8: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,3	84,9	83,0	87,7	89,4	94,6
Taxa de Verdadeiras Positivas [%]	81,8	81,2	82,3	82,6	89,3	90,7
Taxa de Verdadeiras Negativas [%]	96,0	96,4	95,8	97,1	97,3	98,6
Valor Preditivo Negativo [%]	96,6	96,5	96,7	96,8	98,0	98,3
<i>F1 Score</i> [%]	80,8	81,0	80,9	82,6	88,0	90,8
<i>Informedness</i> [%]	77,9	77,7	78,1	79,8	86,7	89,3
<i>Markedness</i> [%]	80,9	81,4	79,7	84,5	87,4	92,9

8.1.3.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 76,7%. A Tabela 8.9 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.9: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em músculos selecionados.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	71,5	79,5	77,2	80,2	85,3	91,7
Taxa de Verdadeiras Positivas [%]	80,2	71,4	75,0	77,2	74,7	81,4
Taxa de Verdadeiras Negativas [%]	90,5	95,4	94,8	95,2	97,3	98,6
Valor Preditivo Negativo [%]	96,0	94,7	95,4	95,7	95,3	96,6
<i>F1 Score</i> [%]	72,8	72,4	73,5	76,0	77,5	83,2
<i>Informedness</i> [%]	70,8	66,9	69,8	72,4	72,0	80,0
<i>Markedness</i> [%]	67,5	74,2	72,6	75,9	80,7	88,4

8.1.4 Utilizando 8 Eletrodos em Volta do Antebraço

Nesta seção, considera-se os sinais de 8 dos 12 eletrodos disponíveis na base de dados NinaPro. Em particular, considera-se apenas os 8 eletrodos igualmente espaçados em volta do antebraço (eletrodos 1 a 8, conforme mostra a Figura 3.1).

8.1.4.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 93,7%. A Tabela 8.10 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.10: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	91,7	94,6	93,7	97,4	98,3	96,6
Taxa de Verdadeiras Positivas [%]	92,5	89,7	91,8	96,3	97,6	94,2
Taxa de Verdadeiras Negativas [%]	97,5	98,6	98,3	99,2	99,5	99,1
Valor Preditivo Negativo [%]	98,6	98,1	98,5	99,3	99,5	98,9
<i>F1 Score</i> [%]	90,2	90,3	91,1	96,1	97,5	94,1
<i>Informedness</i> [%]	90,1	88,4	90,1	95,6	97,1	93,3
<i>Markedness</i> [%]	90,4	92,7	92,2	96,8	97,9	95,5

8.1.4.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 93,6%. A Tabela 8.11 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.11: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	91,7	94,6	93,6	97,4	98,0	96,5
Taxa de Verdadeiras Positivas [%]	92,4	90,0	91,4	96,3	97,3	94,2
Taxa de Verdadeiras Negativas [%]	97,6	98,6	98,3	99,2	99,4	99,1
Valor Preditivo Negativo [%]	98,6	98,1	98,4	99,3	99,5	98,9
<i>F1 Score</i> [%]	90,2	90,5	90,9	96,1	97,0	94,2
<i>Informedness</i> [%]	90,0	88,7	89,7	95,6	96,8	93,4
<i>Markedness</i> [%]	90,4	92,7	92,0	96,7	97,5	95,5

8.1.4.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 88,0%. A Tabela 8.12 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.12: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 8 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	81,9	88,1	87,9	95,3	97,2	93,9
Taxa de Verdadeiras Positivas [%]	93,6	84,4	80,9	91,3	93,8	84,1
Taxa de Verdadeiras Negativas [%]	93,8	97,2	97,4	98,8	99,3	99,0
Valor Preditivo Negativo [%]	98,7	97,1	96,5	98,4	98,8	97,1
<i>F1 Score</i> [%]	85,1	84,3	81,3	91,5	94,5	86,2
<i>Informedness</i> [%]	87,5	81,6	78,3	90,1	93,1	83,2
<i>Markedness</i> [%]	80,7	85,3	84,5	93,7	96,1	91,1

8.1.5 Utilizando 4 Eletrodos em Volta do Antebraço

Nesta seção, considera-se os sinais de 4 dos 12 eletrodos disponíveis na base de dados NinaPro. Em particular, considera-se apenas 4 dos 8 eletrodos igualmente espaçados em volta do antebraço (eletrodos 1, 3, 5 e 7, conforme mostra a Figura 3.1).

8.1.5.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 86,0%. A Tabela 8.13 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.13: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,6	86,5	86,7	92,0	94,8	88,6
Taxa de Verdadeiras Positivas [%]	86,5	81,4	81,1	88,9	93,6	84,3
Taxa de Verdadeiras Negativas [%]	95,7	96,8	96,7	98,1	98,5	97,1
Valor Preditivo Negativo [%]	97,4	96,5	96,4	97,9	98,8	97,1
<i>F1 Score</i> [%]	83,2	81,7	81,2	88,8	92,8	84,0
<i>Informedness</i> [%]	82,3	78,2	77,8	87,0	92,2	81,5
<i>Markedness</i> [%]	82,1	83,1	83,2	90,0	93,6	85,7

8.1.5.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 86,3%. A Tabela 8.14 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.14: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,9	87,2	86,5	92,2	94,6	88,6
Taxa de Verdadeiras Positivas [%]	86,1	82,6	81,3	89,4	94,2	84,0
Taxa de Verdadeiras Negativas [%]	96,0	96,9	96,7	98,0	98,5	97,3
Valor Preditivo Negativo [%]	97,4	96,8	96,5	98,0	98,9	97,0
<i>F1 Score</i> [%]	83,3	82,4	81,5	88,9	93,3	84,0
<i>Informedness</i> [%]	82,1	79,6	78,1	87,4	92,8	81,3
<i>Markedness</i> [%]	82,3	84,0	83,1	90,2	93,6	85,7

8.1.5.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 80,5%. A Tabela 8.15 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.15: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 4 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	74,6	81,0	82,8	89,1	93,1	85,8
Taxa de Verdadeiras Positivas [%]	90,7	76,4	75,4	80,0	87,5	72,8
Taxa de Verdadeiras Negativas [%]	91,2	95,3	96,3	97,8	98,2	97,6
Valor Preditivo Negativo [%]	98,2	95,6	95,4	96,3	97,7	95,0
<i>F1 Score</i> [%]	79,5	76,0	76,4	81,7	88,0	75,7
<i>Informedness</i> [%]	81,9	71,8	71,8	77,8	85,8	70,4
<i>Markedness</i> [%]	72,8	76,6	78,3	85,5	90,9	80,9

8.1.6 Utilizando 2 Eletrodos em Volta do Antebraço

Nesta seção, considera-se os sinais de 2 dos 12 eletrodos disponíveis na base de dados NinaPro. Em particular, considera-se apenas 2 dos 8 eletrodos igualmente espaçados em volta do antebraço (eletrodos 2 e 6, conforme mostra a Figura 3.1).

8.1.6.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.16 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.16: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando utilizando 2 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,9	100,0	99,9	100,0	100,0	100,0
Taxa de Verdadeiras Positivas [%]	99,8	100,0	99,9	100,0	100,0	100,0
Taxa de Verdadeiras Negativas [%]	99,9	100,0	99,9	100,0	100,0	100,0
Valor Preditivo Negativo [%]	99,9	100,0	99,9	100,0	100,0	100,0
<i>F1 Score</i> [%]	99,9	100,0	99,9	100,0	100,0	100,0
<i>Informedness</i> [%]	99,8	100,0	99,8	100,0	100,0	100,0
<i>Markedness</i> [%]	99,9	100,0	99,9	100,0	100,0	100,0

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 76,9 %. A Tabela 8.17 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.17: Desempenho do modelo de regressão logística contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	74,2	76,8	78,3	85,3	86,3	80,7
Taxa de Verdadeiras Positivas [%]	73,3	76,8	74,9	76,4	83,5	76,8
Taxa de Verdadeiras Negativas [%]	93,7	94,5	94,9	96,8	96,4	95,8
Valor Preditivo Negativo [%]	94,9	95,6	95,2	95,6	96,8	95,7
<i>F1 Score</i> [%]	71,9	75,5	75,5	78,0	83,1	76,8
<i>Informedness</i> [%]	67,0	71,4	69,8	73,2	79,9	72,7
<i>Markedness</i> [%]	69,2	72,5	73,6	80,9	83,1	76,4

8.1.6.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 77,9%. A Tabela 8.18 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.18: Desempenho do modelo de rede neural contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	76,4	77,5	79,5	85,8	86,8	81,7
Taxa de Verdadeiras Positivas [%]	74,6	76,4	76,6	78,7	83,2	78,2
Taxa de Verdadeiras Negativas [%]	94,0	94,8	95,2	96,7	96,8	95,9
Valor Preditivo Negativo [%]	95,2	95,5	95,6	96,0	96,8	95,9
<i>F1 Score</i> [%]	73,0	75,9	76,8	79,6	83,3	78,0
<i>Informedness</i> [%]	68,6	71,2	71,9	75,4	80,0	74,1
<i>Markedness</i> [%]	71,6	73,1	75,1	81,8	83,7	77,6

8.1.6.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 70,3%. A Tabela 8.19 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.19: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando utilizando 2 eletrodos em volta do antebraço.

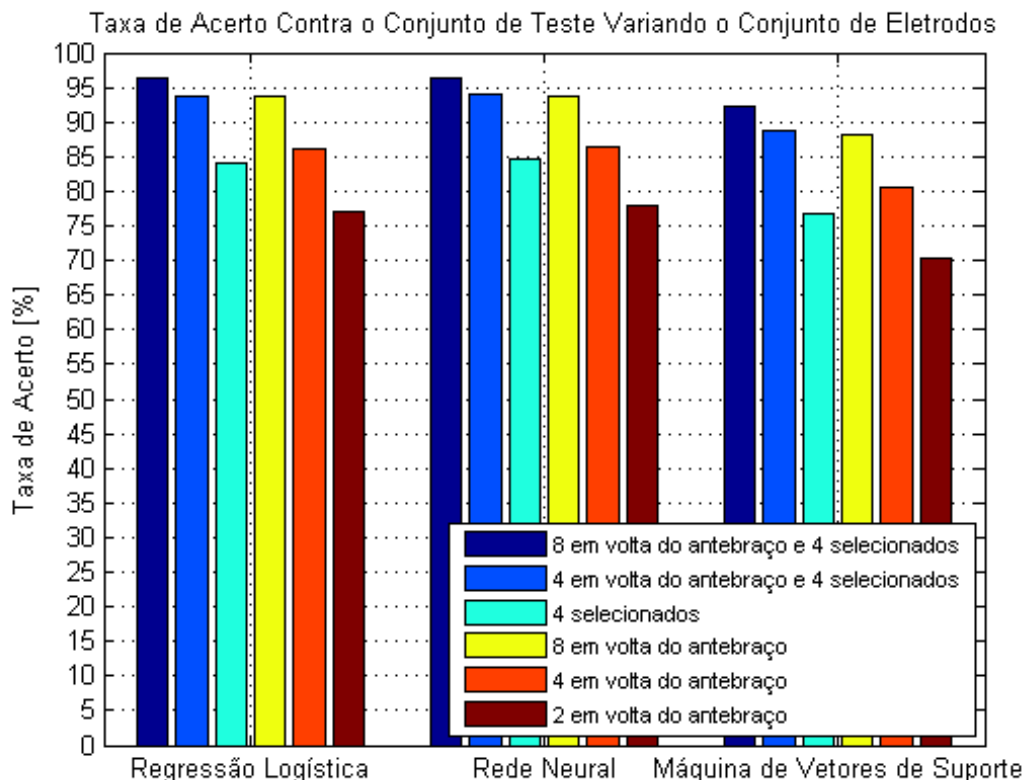
Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	61,5	69,5	72,3	81,5	85,4	82,1
Taxa de Verdadeiras Positivas [%]	80,5	76,2	61,8	68,1	73,9	61,4
Taxa de Verdadeiras Negativas [%]	86,2	91,5	94,9	96,5	97,2	97,9
Valor Preditivo Negativo [%]	95,9	95,4	92,9	94,2	95,2	93,1
<i>F1 Score</i> [%]	67,5	71,1	65,5	71,7	77,4	66,9
<i>Informedness</i> [%]	66,8	67,8	56,7	64,7	71,2	59,4
<i>Markedness</i> [%]	57,5	65,0	65,3	75,7	80,6	75,3

8.1.7 Análise de Desempenho Variando o Conjunto de Eletrodos

A Figura 8.1 mostra a taxa de acerto média de cada um dos modelos de aprendizado para conjuntos de eletrodos diferentes. Como se pode observar, a quantidade de eletrodos utilizados interfere, diretamente, na taxa de acerto média dos modelos de aprendizado. Pela mesma figura, a taxa de acerto dos três modelos parece ser mais influenciada pela quantidade de eletrodos considerados que pela localização deles, no entanto a configuração com apenas 4 eletrodos igualmente espaçados em volta do antebraço possui uma

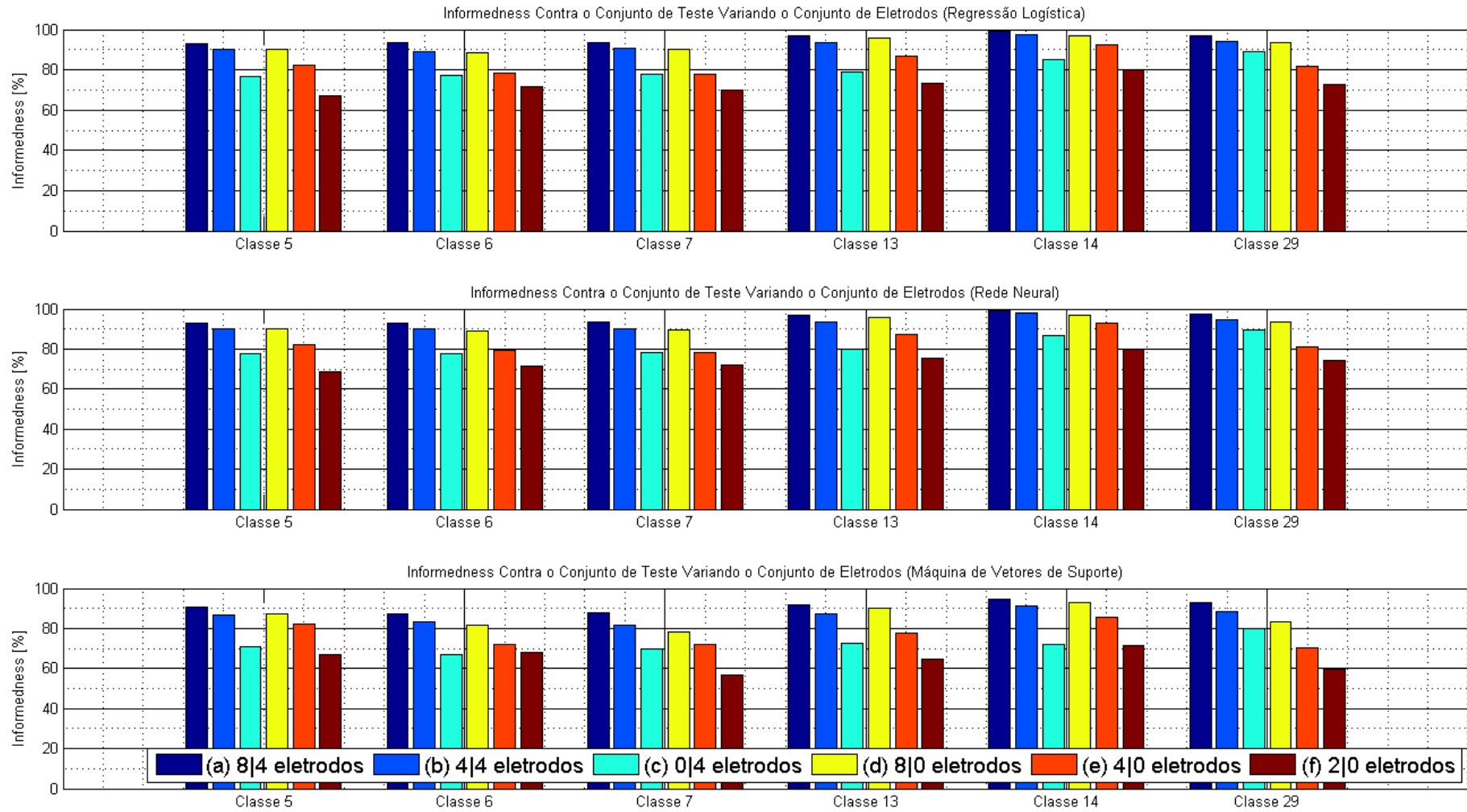
taxa de acerto visivelmente mais alta que a configuração com apenas os 4 eletrodos anatomicamente posicionados em músculos selecionados, mostrando que a localização dos eletrodos também é um fator importante, como pode ser visto na Figura 8.2. Ainda de acordo com a Figura 8.1, nota-se que os modelos de regressão logística e rede neural possuem taxas de acerto bastante semelhantes, enquanto que o modelo de máquina de vetores de suporte apresenta uma taxa de acerto inferior às taxas dos outros modelos em cada um dos casos considerados.

Figura 8.1: Taxa de acerto variando o conjunto de eletrodos.



A Figura 8.2 compara os valores da métrica *informedness* para cada uma das classes, para cada modelo e para cada conjunto de eletrodos. Em (a) são utilizados 8 eletrodos igualmente espaçados em volta do antebraço e 4 eletrodos de posição anatômico precisa; em (b) são utilizados 4 eletrodos igualmente espaçados em volta do antebraço e 4 eletrodos de posição anatômico precisa; em (c) são utilizados 4 eletrodos de posição anatômico precisa; em (d) são utilizados 8 eletrodos igualmente espaçados em volta do antebraço; em (e) são utilizados 4 eletrodos igualmente espaçados em volta do antebraço; em (f) são utilizados 2 eletrodos igualmente espaçados em volta do antebraço. Como observado anteriormente, nesta figura, pode-se notar, por (c) e (e), a importância da localização dos eletrodos para a classificação dos movimentos - em especial, das classes 5, 13, 14 e 29, ou, respectivamente, abdução dos dedos, flexão da mão, extensão da mão e pegada (veja a Figura 3.3). Nota-se que os eletrodos de posição anatômico precisa melhoram, visivelmente, o valor de *informedness* para a classificação da classe 29 (pegada). Provavelmente, isto ocorre porque os eletrodos de posição anatômico precisa incluem os músculos diretamente responsáveis pela flexão e extensão dos dedos - *flexor digitorum superficialis* e *extensor digitorum* (veja a Seção 2.3), respectivamente.

Figura 8.2: *Informedness* variando o conjunto de eletrodos.



8.2 Desempenho Variando o Conjunto de Características

Nesta seção, avalia-se o desempenho dos modelos de aprendizado de máquina para conjuntos variáveis de características. Os seguintes conjuntos são avaliados:

- conjunto de todas as características (Seção 8.2.1);
- conjunto de todas as características, exceto valor RMS (Seção 8.2.2);
- conjunto de todas as características, exceto média aritmética (Seção 8.2.3);
- conjunto de todas as características, exceto variância (Seção 8.2.4);
- conjunto de todas as características, exceto desvio padrão (Seção 8.2.5);
- conjunto de todas as características, exceto obliquidade (Seção 8.2.6);
- conjunto de todas as características, exceto curtose (Seção 8.2.7);
- conjunto de todas as características, exceto centroide espectral (Seção 8.2.8);
- conjunto de todas as características, exceto aquelas que se referem ao domínio do tempo (Seção 8.2.9);
- conjunto de todas as características, exceto aquelas que se referem ao domínio da frequência (Seção 8.2.10);
- somente valor RMS no domínio do tempo (Seção 8.2.11);
- somente média aritmética no domínio do tempo (Seção 8.2.12);
- somente variância no domínio do tempo (Seção 8.2.13);
- somente desvio padrão no domínio do tempo (Seção 8.2.14);
- somente obliquidade no domínio do tempo (Seção 8.2.15);
- somente curtose no domínio do tempo (Seção 8.2.16);
- somente valor RMS no domínio da frequência (Seção 8.2.17);
- somente média aritmética no domínio da frequência (Seção 8.2.18);
- somente variância no domínio da frequência (Seção 8.2.19);
- somente desvio padrão no domínio da frequência (Seção 8.2.20);
- somente obliquidade no domínio da frequência (Seção 8.2.21);
- somente curtose no domínio da frequência (Seção 8.2.22);
- somente centroide espectral no domínio da frequência (Seção 8.2.23);

Após a exibição de resultados, na Seção 8.2.24, faz-se uma avaliação de desempenho geral, comparando todos os conjuntos considerados.

8.2.1 Utilizando Todas as Características

Nesta seção, considera-se todas as características descritas na Seção 4.5.

8.2.1.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,3%. A Tabela 8.20 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.20: Desempenho do modelo de regressão logística contra o conjunto de teste quando considerando todas as características.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,3	97,8	95,1	98,6	99,3	98,7
Taxa de Verdadeiras Positivas [%]	94,8	93,9	95,0	97,5	99,1	97,3
Taxa de Verdadeiras Negativas [%]	98,3	99,4	98,6	99,6	99,8	99,6
Valor Preditivo Negativo [%]	99,0	98,8	99,0	99,5	99,8	99,5
<i>F1 Score</i> [%]	93,4	94,7	93,8	97,5	99,0	97,5
<i>Informedness</i> [%]	93,1	93,3	93,6	97,1	98,9	97,0
<i>Markedness</i> [%]	93,4	96,7	94,2	98,2	99,1	98,3

8.2.1.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,3%. A Tabela 8.21 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.21: Desempenho do modelo de rede neural contra o conjunto de teste quando considerando todas as características.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,8	97,6	94,8	98,7	99,1	98,9
Taxa de Verdadeiras Positivas [%]	94,6	93,6	95,0	97,4	99,2	97,9
Taxa de Verdadeiras Negativas [%]	98,5	99,4	98,5	99,6	99,7	99,7
Valor Preditivo Negativo [%]	99,0	98,8	99,1	99,5	99,8	99,6
<i>F1 Score</i> [%]	93,5	94,6	93,7	97,5	98,9	97,9
<i>Informedness</i> [%]	93,1	93,1	93,5	97,0	98,9	97,6
<i>Markedness</i> [%]	93,8	96,4	93,8	98,2	98,9	98,5

8.2.1.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 92,2 %. A Tabela 8.22 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.22: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando considerando todas as características.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	86,8	92,5	92,5	96,7	98,1	98,1
Taxa de Verdadeiras Positivas [%]	94,9	88,9	89,5	92,4	94,9	92,9
Taxa de Verdadeiras Negativas [%]	95,5	98,3	98,1	99,2	99,6	99,7
Valor Preditivo Negativo [%]	99,0	97,9	98,1	98,6	99,0	98,7
<i>F1 Score</i> [%]	88,8	88,9	88,8	93,1	95,5	94,1
<i>Informedness</i> [%]	90,4	87,2	87,7	91,7	94,4	92,6
<i>Markedness</i> [%]	85,9	90,5	90,6	95,3	97,2	96,9

8.2.2 Ignorando o Valor RMS

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto o valor RMS do sinal (nos domínios do tempo e da frequência), que foi descrito na Seção 4.5.1.

8.2.2.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,8 %. A Tabela 8.23 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.23: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o valor RMS do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,2	98,2	95,6	98,9	99,4	98,9
Taxa de Verdadeiras Positivas [%]	95,9	94,7	95,4	97,5	99,0	98,0
Taxa de Verdadeiras Negativas [%]	98,5	99,5	98,7	99,7	99,8	99,7
Valor Preditivo Negativo [%]	99,2	99,0	99,1	99,5	99,8	99,6
<i>F1 Score</i> [%]	94,6	95,5	94,4	97,7	99,1	98,1
<i>Informedness</i> [%]	94,5	94,3	94,1	97,2	98,9	97,7
<i>Markedness</i> [%]	94,5	97,2	94,8	98,5	99,2	98,6

8.2.2.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,9 %. A Tabela 8.24 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.24: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o valor RMS do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	96,0	98,3	95,6	98,7	99,3	98,9
Taxa de Verdadeiras Positivas [%]	95,5	94,8	95,6	97,8	99,3	98,5
Taxa de Verdadeiras Negativas [%]	98,8	99,6	98,7	99,6	99,8	99,6
Valor Preditivo Negativo [%]	99,1	99,0	99,2	99,6	99,8	99,7
<i>F1 Score</i> [%]	94,7	95,7	94,5	97,8	99,1	98,3
<i>Informedness</i> [%]	94,3	94,4	94,4	97,5	99,1	98,2
<i>Markedness</i> [%]	95,2	97,4	94,8	98,3	99,1	98,6

8.2.2.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 93,0%. A Tabela 8.25 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.25: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o valor RMS do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	88,5	93,0	93,0	96,9	98,9	98,3
Taxa de Verdadeiras Positivas [%]	96,0	90,6	89,1	93,5	95,9	93,1
Taxa de Verdadeiras Negativas [%]	96,1	98,3	98,3	99,3	99,7	99,8
Valor Preditivo Negativo [%]	99,2	98,2	98,0	98,8	99,2	98,7
<i>F1 Score</i> [%]	90,4	90,2	88,9	94,2	96,7	94,3
<i>Informedness</i> [%]	92,1	88,9	87,4	92,9	95,7	92,9
<i>Markedness</i> [%]	87,8	91,3	91,0	95,7	98,2	97,0

8.2.3 Ignorando a Média Aritmética

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto a média aritmética do sinal (nos domínios do tempo e da frequência), que foi descrita na Seção 4.5.2.

8.2.3.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,6%. A Tabela 8.26 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.26: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a média aritmética do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,2	98,0	95,4	98,8	99,3	98,8
Taxa de Verdadeiras Positivas [%]	95,4	94,1	95,9	97,4	99,1	97,9
Taxa de Verdadeiras Negativas [%]	98,5	99,5	98,7	99,6	99,8	99,6
Valor Preditivo Negativo [%]	99,1	98,9	99,2	99,5	99,8	99,6
<i>F1 Score</i> [%]	94,2	95,0	94,7	97,6	99,0	97,9
<i>Informedness</i> [%]	93,9	93,6	94,6	97,0	98,9	97,5
<i>Markedness</i> [%]	94,3	97,0	94,7	98,3	99,2	98,4

8.2.3.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,7%. A Tabela 8.27 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.27: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a média aritmética do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,5	98,2	95,6	98,6	99,1	98,8
Taxa de Verdadeiras Positivas [%]	95,0	94,5	95,7	97,7	99,2	98,1
Taxa de Verdadeiras Negativas [%]	98,7	99,5	98,7	99,6	99,7	99,6
Valor Preditivo Negativo [%]	99,0	99,0	99,2	99,5	99,8	99,6
<i>F1 Score</i> [%]	94,2	95,4	94,6	97,7	99,0	98,0
<i>Informedness</i> [%]	93,7	94,0	94,4	97,3	99,0	97,8
<i>Markedness</i> [%]	94,6	97,2	94,8	98,2	99,0	98,5

8.2.3.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 93,0%. A Tabela 8.28 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.28: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a média aritmética do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	88,8	92,8	93,0	97,3	98,0	98,8
Taxa de Verdadeiras Positivas [%]	95,0	89,9	88,9	94,2	95,8	94,6
Taxa de Verdadeiras Negativas [%]	96,4	98,2	98,3	99,3	99,5	99,8
Valor Preditivo Negativo [%]	99,0	98,1	97,9	98,9	99,2	99,0
<i>F1 Score</i> [%]	90,2	89,6	88,9	94,7	96,2	95,6
<i>Informedness</i> [%]	91,4	88,1	87,2	93,5	95,3	94,5
<i>Markedness</i> [%]	87,8	91,0	91,0	96,2	97,2	97,8

8.2.4 Ignorando a Variância

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto a variância do sinal (nos domínios do tempo e da frequência), que foi descrita na Seção 4.5.3.

8.2.4.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,9%. A Tabela 8.29 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.29: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a variância do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,5	98,2	95,8	99,0	99,4	99,0
Taxa de Verdadeiras Positivas [%]	95,9	94,9	95,5	97,5	99,2	98,4
Taxa de Verdadeiras Negativas [%]	98,6	99,5	98,8	99,7	99,8	99,7
Valor Preditivo Negativo [%]	99,2	99,0	99,1	99,5	99,8	99,7
<i>F1 Score</i> [%]	94,6	95,6	94,6	97,8	99,2	98,4
<i>Informedness</i> [%]	94,5	94,4	94,3	97,3	99,1	98,1
<i>Markedness</i> [%]	94,8	97,3	95,0	98,6	99,3	98,8

8.2.4.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,8%. A Tabela 8.30 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.30: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a variância do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,6	98,0	95,2	98,9	99,4	98,9
Taxa de Verdadeiras Positivas [%]	95,3	94,7	95,6	97,3	99,3	98,4
Taxa de Verdadeiras Negativas [%]	98,7	99,5	98,6	99,7	99,8	99,6
Valor Preditivo Negativo [%]	99,1	99,0	99,2	99,5	99,8	99,7
<i>F1 Score</i> [%]	94,4	95,6	94,3	97,6	99,2	98,3
<i>Informedness</i> [%]	94,0	94,3	94,2	97,1	99,2	98,1
<i>Markedness</i> [%]	94,8	97,0	94,4	98,4	99,3	98,6

8.2.4.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 93,2%. A Tabela 8.31 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.31: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a variância do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	88,6	92,9	92,9	97,5	98,7	98,8
Taxa de Verdadeiras Positivas [%]	96,2	90,6	89,1	93,3	96,2	94,1
Taxa de Verdadeiras Negativas [%]	96,2	98,3	98,3	99,5	99,7	99,9
Valor Preditivo Negativo [%]	99,2	98,3	98,0	98,8	99,3	98,9
<i>F1 Score</i> [%]	90,7	90,0	88,8	94,2	96,8	95,3
<i>Informedness</i> [%]	92,4	88,9	87,4	92,8	95,9	94,0
<i>Markedness</i> [%]	87,9	91,2	90,9	96,3	98,0	97,7

8.2.5 Ignorando o Desvio Padrão

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto o desvio padrão do sinal (nos domínios do tempo e da frequência), que foi descrito na Seção 4.5.4.

8.2.5.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,8%. A Tabela 8.32 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.32: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o desvio padrão do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,3	98,0	95,6	98,9	99,4	99,1
Taxa de Verdadeiras Positivas [%]	95,8	94,8	95,3	97,5	99,1	98,2
Taxa de Verdadeiras Negativas [%]	98,6	99,5	98,7	99,7	99,8	99,7
Valor Preditivo Negativo [%]	99,2	99,0	99,1	99,5	99,8	99,6
<i>F1 Score</i> [%]	94,5	95,5	94,3	97,7	99,1	98,3
<i>Informedness</i> [%]	94,4	94,3	94,0	97,2	98,9	98,0
<i>Markedness</i> [%]	94,5	97,1	94,7	98,5	99,3	98,8

8.2.5.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,9%. A Tabela 8.33 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.33: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o desvio padrão do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	95,9	98,4	95,6	99,0	99,4	98,8
Taxa de Verdadeiras Positivas [%]	95,6	95,0	96,0	97,4	99,3	98,4
Taxa de Verdadeiras Negativas [%]	98,7	99,6	98,7	99,7	99,8	99,6
Valor Preditivo Negativo [%]	99,1	99,1	99,2	99,5	99,8	99,7
<i>F1 Score</i> [%]	94,7	95,8	94,8	97,7	99,2	98,2
<i>Informedness</i> [%]	94,3	94,6	94,8	97,1	99,1	98,1
<i>Markedness</i> [%]	95,1	97,5	94,9	98,6	99,2	98,5

8.2.5.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 92,8%. A Tabela 8.34 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.34: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o desvio padrão do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	88,5	92,9	93,0	96,5	98,6	98,2
Taxa de Verdadeiras Positivas [%]	96,1	90,6	88,6	92,7	95,5	93,4
Taxa de Verdadeiras Negativas [%]	96,0	98,2	98,3	99,2	99,7	99,8
Valor Preditivo Negativo [%]	99,2	98,2	97,9	98,6	99,2	98,8
<i>F1 Score</i> [%]	90,4	90,0	88,5	93,4	96,3	94,5
<i>Informedness</i> [%]	92,1	88,8	86,9	92,0	95,3	93,2
<i>Markedness</i> [%]	87,8	91,2	90,9	95,2	97,8	97,1

8.2.6 Ignorando a Obliquidade

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto a obliquidade do sinal (nos domínios do tempo e da frequência), que foi descrita na Seção 4.5.5.

8.2.6.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 94,3%. A Tabela 8.35 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.35: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a obliquidade do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,1	95,1	93,1	97,6	98,7	98,2
Taxa de Verdadeiras Positivas [%]	92,7	90,1	93,3	95,3	98,5	96,1
Taxa de Verdadeiras Negativas [%]	97,6	98,8	98,1	99,3	99,6	99,5
Valor Preditivo Negativo [%]	98,6	98,2	98,7	99,1	99,7	99,3
<i>F1 Score</i> [%]	90,8	90,9	91,7	95,6	98,3	96,3
<i>Informedness</i> [%]	90,4	89,0	91,4	94,7	98,2	95,6
<i>Markedness</i> [%]	90,8	93,3	91,9	96,8	98,5	97,5

8.2.6.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 94,4%. A Tabela 8.36 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.36: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a obliquidade do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,6	95,5	93,1	97,7	98,6	97,7
Taxa de Verdadeiras Positivas [%]	92,1	91,1	92,9	95,7	98,5	96,2
Taxa de Verdadeiras Negativas [%]	97,8	98,9	98,1	99,3	99,6	99,3
Valor Preditivo Negativo [%]	98,5	98,3	98,7	99,2	99,7	99,3
<i>F1 Score</i> [%]	90,7	91,8	91,4	95,9	98,2	96,1
<i>Informedness</i> [%]	90,0	90,1	91,1	95,1	98,1	95,6
<i>Markedness</i> [%]	91,2	93,9	91,8	96,9	98,3	97,1

8.2.6.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 88,4%. A Tabela 8.37 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.37: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a obliquidade do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	80,9	89,3	89,4	94,5	96,4	95,7
Taxa de Verdadeiras Positivas [%]	91,0	85,0	84,8	88,6	93,3	87,5
Taxa de Verdadeiras Negativas [%]	93,5	97,6	97,5	98,7	99,1	99,4
Valor Preditivo Negativo [%]	98,2	97,2	97,2	97,9	98,8	97,8
<i>F1 Score</i> [%]	83,3	85,4	84,1	89,5	93,8	89,1
<i>Informedness</i> [%]	84,6	82,6	82,4	87,3	92,5	87,0
<i>Markedness</i> [%]	79,2	86,6	86,7	92,4	95,2	93,5

8.2.7 Ignorando a Curtose

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto a curtose do sinal (nos domínios do tempo e da frequência), que foi descrita na Seção 4.5.6.

8.2.7.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 95,4%. A Tabela 8.38 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.38: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando a curtose do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,4	96,4	94,0	98,3	99,0	98,5
Taxa de Verdadeiras Positivas [%]	93,5	92,6	94,4	96,4	98,8	96,5
Taxa de Verdadeiras Negativas [%]	98,1	99,1	98,3	99,5	99,7	99,6
Valor Preditivo Negativo [%]	98,8	98,6	98,9	99,3	99,7	99,3
<i>F1 Score</i> [%]	92,1	93,2	92,9	96,7	98,6	96,8
<i>Informedness</i> [%]	91,6	91,7	92,8	95,9	98,5	96,2
<i>Markedness</i> [%]	92,2	95,0	93,0	97,6	98,7	97,9

8.2.7.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 95,4%. A Tabela 8.39 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.39: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando a curtose do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,9	96,2	94,1	98,2	98,9	98,4
Taxa de Verdadeiras Positivas [%]	92,9	93,3	94,7	95,9	98,8	97,0
Taxa de Verdadeiras Negativas [%]	98,2	99,1	98,3	99,5	99,6	99,5
Valor Preditivo Negativo [%]	98,6	98,7	99,0	99,2	99,7	99,4
<i>F1 Score</i> [%]	91,9	93,7	93,1	96,3	98,6	97,1
<i>Informedness</i> [%]	91,1	92,4	93,0	95,5	98,5	96,6
<i>Markedness</i> [%]	92,6	95,0	93,1	97,5	98,7	97,9

8.2.7.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 91,1%. A Tabela 8.40 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.40: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando a curtose do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,0	92,2	92,7	95,6	97,9	97,4
Taxa de Verdadeiras Positivas [%]	94,0	88,2	87,6	90,9	94,2	91,8
Taxa de Verdadeiras Negativas [%]	94,6	98,2	98,3	99,0	99,5	99,5
Valor Preditivo Negativo [%]	98,8	97,8	97,7	98,3	98,9	98,5
<i>F1 Score</i> [%]	86,6	88,5	87,9	91,7	95,0	93,0
<i>Informedness</i> [%]	88,6	86,4	86,0	90,0	93,7	91,4
<i>Markedness</i> [%]	82,8	90,1	90,5	94,0	96,8	96,0

8.2.8 Ignorando o Centroide Espectral

Nesta seção, considera-se todas as características descritas na Seção 4.5 exceto o centroide espectral do sinal (no domínio da frequência), que foi descrito na Seção 4.5.7.

8.2.8.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 94,9%. A Tabela 8.41 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.41: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando o centroide espectral do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,9	96,5	93,7	97,8	98,5	98,2
Taxa de Verdadeiras Positivas [%]	93,1	91,8	93,6	96,4	97,9	96,9
Taxa de Verdadeiras Negativas [%]	97,9	99,1	98,3	99,4	99,6	99,5
Valor Preditivo Negativo [%]	98,7	98,5	98,8	99,3	99,6	99,4
<i>F1 Score</i> [%]	91,5	92,9	92,2	96,3	97,8	96,8
<i>Informedness</i> [%]	91,0	91,0	92,0	95,8	97,5	96,4
<i>Markedness</i> [%]	91,6	95,0	92,6	97,2	98,1	97,6

8.2.8.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 94,9%. A Tabela 8.42 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.42: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando o centroide espectral do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,5	96,4	93,6	97,7	98,3	98,2
Taxa de Verdadeiras Positivas [%]	92,6	91,8	94,1	96,3	98,3	96,6
Taxa de Verdadeiras Negativas [%]	98,1	99,1	98,2	99,4	99,5	99,5
Valor Preditivo Negativo [%]	98,6	98,5	98,9	99,3	99,7	99,4
<i>F1 Score</i> [%]	91,5	92,8	92,5	96,3	97,9	96,7
<i>Informedness</i> [%]	90,7	90,9	92,3	95,7	97,9	96,1
<i>Markedness</i> [%]	92,1	94,9	92,5	97,0	98,0	97,6

8.2.8.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 89,6%. A Tabela 8.43 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.43: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando o centroide espectral do sinal.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	82,9	90,3	91,2	94,8	96,8	96,3
Taxa de Verdadeiras Positivas [%]	92,7	85,5	86,8	89,9	92,7	89,7
Taxa de Verdadeiras Negativas [%]	94,2	97,8	97,8	98,8	99,2	99,5
Valor Preditivo Negativo [%]	98,6	97,3	97,6	98,1	98,6	98,2
<i>F1 Score</i> [%]	85,2	85,6	86,4	90,5	93,4	90,8
<i>Informedness</i> [%]	87,0	83,4	84,7	88,8	92,0	89,3
<i>Markedness</i> [%]	81,6	87,7	88,8	93,0	95,5	94,5

8.2.9 Ignorando o Domínio do Tempo

Nesta seção, considera-se apenas as características descritas na Seção 4.5 que são medidas no domínio da frequência.

8.2.9.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 95,7%. A Tabela 8.44 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.44: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,2	97,1	95,5	97,6	99,3	98,8
Taxa de Verdadeiras Positivas [%]	95,4	94,0	94,5	95,1	98,7	96,7
Taxa de Verdadeiras Negativas [%]	97,9	99,2	98,8	99,4	99,8	99,6
Valor Preditivo Negativo [%]	99,1	98,9	99,0	99,1	99,7	99,4
<i>F1 Score</i> [%]	93,1	94,4	93,8	95,4	98,8	97,1
<i>Informedness</i> [%]	93,4	93,3	93,3	94,5	98,5	96,4
<i>Markedness</i> [%]	92,4	96,0	94,5	96,7	99,1	98,2

8.2.9.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 95,6%. A Tabela 8.45 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.45: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,6	97,1	94,8	97,9	99,2	98,5
Taxa de Verdadeiras Positivas [%]	95,1	93,5	94,4	94,9	99,1	96,8
Taxa de Verdadeiras Negativas [%]	98,0	99,2	98,5	99,5	99,7	99,6
Valor Preditivo Negativo [%]	99,1	98,8	99,0	99,0	99,8	99,4
<i>F1 Score</i> [%]	93,1	94,0	93,3	95,4	98,9	97,0
<i>Informedness</i> [%]	93,1	92,7	93,0	94,4	98,9	96,5
<i>Markedness</i> [%]	92,7	95,9	93,8	97,0	99,0	98,0

8.2.9.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 92,4%. A Tabela 8.46 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.46: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando todas as características do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	86,1	94,3	93,6	95,3	98,7	98,1
Taxa de Verdadeiras Positivas [%]	95,8	90,5	89,9	90,7	95,0	92,2
Taxa de Verdadeiras Negativas [%]	95,4	98,7	98,3	98,9	99,7	99,7
Valor Preditivo Negativo [%]	99,2	98,3	98,1	98,3	99,1	98,6
<i>F1 Score</i> [%]	89,0	90,6	89,7	91,3	95,9	93,6
<i>Informedness</i> [%]	91,3	89,2	88,2	89,7	94,7	91,9
<i>Markedness</i> [%]	85,4	92,6	91,7	93,6	97,8	96,7

8.2.10 Ignorando o Domínio da Frequência

Nesta seção, considera-se apenas as características descritas na Seção 4.5 que são medidas no domínio do tempo.

8.2.10.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 82,8%. A Tabela 8.47 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.47: Desempenho do modelo de regressão logística contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	80,7	87,0	85,4	87,1	88,3	88,2
Taxa de Verdadeiras Positivas [%]	80,6	80,6	82,1	83,6	85,1	85,0
Taxa de Verdadeiras Negativas [%]	94,8	96,8	96,3	96,9	97,1	97,2
Valor Preditivo Negativo [%]	96,3	96,4	96,7	96,9	97,2	97,2
<i>F1 Score</i> [%]	77,9	81,2	81,3	83,1	84,8	84,3
<i>Informedness</i> [%]	75,5	77,5	78,5	80,6	82,2	82,2
<i>Markedness</i> [%]	77,1	83,4	82,1	84,1	85,5	85,5

8.2.10.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 83,3%. A Tabela 8.48 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.48: Desempenho do modelo de rede neural contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	81,3	87,6	85,1	88,9	88,3	88,2
Taxa de Verdadeiras Positivas [%]	80,7	81,0	83,1	83,4	85,9	85,9
Taxa de Verdadeiras Negativas [%]	95,0	97,0	96,2	97,3	97,0	97,2
Valor Preditivo Negativo [%]	96,3	96,5	96,8	96,9	97,3	97,4
<i>F1 Score</i> [%]	78,5	81,8	81,8	83,6	85,3	84,9
<i>Informedness</i> [%]	75,8	78,1	79,3	80,8	82,9	83,1
<i>Markedness</i> [%]	77,6	84,1	82,0	85,8	85,6	85,6

8.2.10.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 72,2 %. A Tabela 8.49 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.49: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando ignorando todas as características do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	66,1	73,4	79,8	81,3	80,8	81,9
Taxa de Verdadeiras Positivas [%]	79,3	69,1	74,4	70,8	72,2	67,3
Taxa de Verdadeiras Negativas [%]	89,0	93,9	94,9	95,8	95,8	97,1
Valor Preditivo Negativo [%]	95,9	94,2	95,2	94,6	94,9	94,2
<i>F1 Score</i> [%]	69,6	68,3	73,7	71,4	73,4	69,0
<i>Informedness</i> [%]	68,4	63,0	69,3	66,6	68,0	64,4
<i>Markedness</i> [%]	62,1	67,7	75,0	76,0	75,7	76,1

8.2.11 Somente o Valor RMS no Domínio do Tempo

Nesta seção, considera-se apenas o valor RMS do sinal, que foi descrito na Seção 4.5.1, no domínio do tempo.

8.2.11.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.50 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.50: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o valor RMS do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,9	99,9	99,9	99,8	99,9	99,9
Taxa de Verdadeiras Positivas [%]	99,8	99,8	100,0	99,9	99,9	99,9
Taxa de Verdadeiras Negativas [%]	99,9	99,9	99,9	99,9	99,9	99,9
Valor Preditivo Negativo [%]	99,9	99,9	100,0	100,0	100,0	99,9
<i>F1 Score</i> [%]	99,9	99,9	99,9	99,9	99,9	99,9
<i>Informedness</i> [%]	99,8	99,8	99,9	99,9	99,9	99,9
<i>Markedness</i> [%]	99,9	99,9	99,9	99,8	99,9	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 57,5 %. A Tabela 8.51 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.51: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	59,0	57,9	66,4	62,7	60,4	65,1
Taxa de Verdadeiras Positivas [%]	53,5	53,0	59,2	59,1	59,1	61,3
Taxa de Verdadeiras Negativas [%]	91,2	90,6	92,9	91,4	90,5	92,2
Valor Preditivo Negativo [%]	91,1	91,0	92,3	92,2	92,1	92,6
<i>F1 Score</i> [%]	57,3	56,6	63,0	60,6	60,8	63,1
<i>Informedness</i> [%]	44,8	43,7	52,2	50,6	49,6	53,5
<i>Markedness</i> [%]	50,2	48,9	58,8	54,9	52,5	57,8

8.2.11.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 57,9 %. A Tabela 8.52 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.52: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	59,1	58,1	66,6	62,5	61,9	65,2
Taxa de Verdadeiras Positivas [%]	53,0	53,8	58,8	59,4	60,8	61,4
Taxa de Verdadeiras Negativas [%]	91,3	90,5	93,0	91,2	91,0	92,3
Valor Preditivo Negativo [%]	91,0	91,1	92,2	92,2	92,4	92,7
<i>F1 Score</i> [%]	56,6	56,8	62,3	60,6	62,5	63,2
<i>Informedness</i> [%]	44,3	44,4	51,8	50,7	51,8	53,7
<i>Markedness</i> [%]	50,2	49,3	58,8	54,7	54,3	57,9

8.2.11.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 52,2 %. A Tabela 8.53 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.53: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	49,0	49,3	62,0	62,2	59,9	68,0
Taxa de Verdadeiras Positivas [%]	71,2	55,1	54,5	53,6	42,9	36,0
Taxa de Verdadeiras Negativas [%]	80,7	86,3	92,3	91,4	94,2	97,6
Valor Preditivo Negativo [%]	93,7	91,0	91,5	91,2	89,6	88,7
<i>F1 Score</i> [%]	58,0	55,2	57,7	55,7	52,7	44,6
<i>Informedness</i> [%]	51,9	41,4	46,8	45,0	37,1	33,6
<i>Markedness</i> [%]	42,7	40,3	53,5	53,4	49,5	56,8

8.2.12 Somente a Média Aritmética no Domínio do Tempo

Nesta seção, considera-se apenas a média aritmética do sinal, que foi descrita na Seção 4.5.2, no domínio do tempo.

8.2.12.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 96,5 %. A Tabela 8.54 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.54: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a média aritmética do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	96,1	96,3	96,7	97,1	97,9	97,8
Taxa de Verdadeiras Positivas [%]	95,5	95,4	95,0	98,0	97,5	97,7
Taxa de Verdadeiras Negativas [%]	99,0	99,1	99,2	99,3	99,5	99,5
Valor Preditivo Negativo [%]	99,1	99,1	99,0	99,6	99,5	99,5
<i>F1 Score</i> [%]	95,4	95,4	95,4	97,3	97,5	97,6
<i>Informedness</i> [%]	94,5	94,5	94,3	97,3	97,0	97,2
<i>Markedness</i> [%]	95,2	95,4	95,7	96,8	97,4	97,4

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 17,1 %. A Tabela 8.55 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.55: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	19,9	19,7	20,6	24,2	24,8	26,0
Taxa de Verdadeiras Positivas [%]	19,4	15,2	19,2	15,7	15,1	17,8
Taxa de Verdadeiras Negativas [%]	80,1	82,2	81,0	85,3	85,9	85,7
Valor Preditivo Negativo [%]	83,2	82,8	83,4	83,5	83,5	84,0
<i>F1 Score</i> [%]	42,5	41,9	42,9	39,9	39,7	40,5
<i>Informedness</i> [%]	-0,4	-2,4	0,3	1,1	1,1	3,5
<i>Markedness</i> [%]	3,2	2,5	4,1	7,7	8,3	10,0

8.2.12.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 17,5%. A Tabela 8.56 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.56: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	21,0	19,0	21,8	24,8	25,6	25,8
Taxa de Verdadeiras Positivas [%]	21,1	15,5	21,1	14,9	16,4	16,4
Taxa de Verdadeiras Negativas [%]	80,4	81,5	80,8	86,1	86,1	86,0
Valor Preditivo Negativo [%]	83,6	82,7	83,7	83,5	83,7	83,7
<i>F1 Score</i> [%]	43,3	42,1	43,1	39,8	40,0	40,8
<i>Informedness</i> [%]	1,5	-2,9	2,0	1,1	2,5	2,4
<i>Markedness</i> [%]	4,7	1,8	5,5	8,3	9,4	9,5

8.2.12.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 17,6%. A Tabela 8.57 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.57: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	19,3	19,8	22,6	32,1	36,5	41,2
Taxa de Verdadeiras Positivas [%]	52,5	18,0	15,7	7,5	5,9	6,3
Taxa de Verdadeiras Negativas [%]	55,2	79,5	82,9	92,6	94,5	96,2
Valor Preditivo Negativo [%]	85,9	82,8	83,0	83,3	83,3	83,7
<i>F1 Score</i> [%]	39,2	40,5	38,7	28,0	22,4	20,3
<i>Informedness</i> [%]	7,8	-2,4	-1,2	0,2	0,4	2,6
<i>Markedness</i> [%]	5,2	2,6	5,7	15,5	19,9	25,0

8.2.13 Somente a Variância no Domínio do Tempo

Nesta seção, considera-se apenas a variância do sinal, que foi descrita na Seção 4.5.3, no domínio do tempo.

8.2.13.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,8 %. A Tabela 8.58 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.58: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a variância do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,9	99,9	99,9	99,7	99,9	99,9
Taxa de Verdadeiras Positivas [%]	99,8	99,7	100,0	99,9	99,9	99,9
Taxa de Verdadeiras Negativas [%]	99,9	99,9	99,9	99,9	99,9	99,9
Valor Preditivo Negativo [%]	99,9	99,9	100,0	99,9	99,9	99,9
<i>F1 Score</i> [%]	99,8	99,8	99,9	99,8	99,9	99,9
<i>Informedness</i> [%]	99,7	99,7	99,9	99,8	99,9	99,9
<i>Markedness</i> [%]	99,9	99,8	99,9	99,7	99,9	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 57,9 %. A Tabela 8.59 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.59: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	59,8	57,3	67,1	62,2	62,0	65,6
Taxa de Verdadeiras Positivas [%]	54,4	51,9	59,4	58,8	60,6	62,5
Taxa de Verdadeiras Negativas [%]	91,2	90,8	93,1	91,3	90,7	92,1
Valor Preditivo Negativo [%]	91,3	90,8	92,3	92,1	92,3	92,9
<i>F1 Score</i> [%]	58,1	55,8	63,3	60,4	62,3	63,6
<i>Informedness</i> [%]	45,7	42,8	52,5	50,2	51,4	54,7
<i>Markedness</i> [%]	51,1	48,2	59,4	54,3	54,4	58,5

8.2.13.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 58,0%. A Tabela 8.60 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.60: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	59,0	56,9	67,6	62,8	61,7	65,1
Taxa de Verdadeiras Positivas [%]	53,0	52,2	60,4	59,7	61,3	61,5
Taxa de Verdadeiras Negativas [%]	91,1	90,8	93,1	91,4	90,7	92,3
Valor Preditivo Negativo [%]	91,0	90,9	92,5	92,3	92,5	92,7
<i>F1 Score</i> [%]	56,7	55,8	63,2	61,2	62,5	63,2
<i>Informedness</i> [%]	44,2	43,0	53,5	51,1	52,0	53,9
<i>Markedness</i> [%]	50,0	47,8	60,1	55,1	54,2	57,8

8.2.13.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 52,1%. A Tabela 8.61 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.61: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a variância do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	49,7	51,2	61,3	59,5	61,0	66,6
Taxa de Verdadeiras Positivas [%]	72,1	54,4	55,0	52,3	43,5	35,2
Taxa de Verdadeiras Negativas [%]	80,9	86,8	92,0	90,9	94,3	97,4
Valor Preditivo Negativo [%]	93,9	90,9	91,5	90,9	89,7	88,6
<i>F1 Score</i> [%]	58,4	55,4	58,7	54,9	52,0	44,0
<i>Informedness</i> [%]	53,0	41,2	47,0	43,2	37,8	32,7
<i>Markedness</i> [%]	43,6	42,1	52,8	50,5	50,7	55,3

8.2.14 Somente o Desvio Padrão no Domínio do Tempo

Nesta seção, considera-se apenas o desvio padrão do sinal, que foi descrito na Seção 4.5.4, no domínio do tempo.

8.2.14.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9%. A Tabela 8.62 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.62: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o desvio padrão do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,9	99,9	99,9	99,8	99,9	99,9
Taxa de Verdadeiras Positivas [%]	99,9	99,8	100,0	99,9	99,9	99,9
Taxa de Verdadeiras Negativas [%]	99,9	99,9	99,9	99,9	99,9	100,0
Valor Preditivo Negativo [%]	99,9	99,9	100,0	99,9	100,0	99,9
<i>F1 Score</i> [%]	99,9	99,9	99,9	99,9	99,9	99,9
<i>Informedness</i> [%]	99,9	99,8	99,9	99,9	99,9	99,9
<i>Markedness</i> [%]	99,9	99,9	99,9	99,8	99,9	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 57,5 %. A Tabela 8.63 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.63: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	59,8	57,1	66,0	62,5	61,0	64,7
Taxa de Verdadeiras Positivas [%]	53,6	53,0	58,3	58,7	60,3	61,4
Taxa de Verdadeiras Negativas [%]	91,3	90,4	92,9	91,4	90,7	92,1
Valor Preditivo Negativo [%]	91,1	90,9	92,1	92,1	92,3	92,7
<i>F1 Score</i> [%]	57,0	56,3	62,4	60,7	62,3	63,4
<i>Informedness</i> [%]	44,9	43,4	51,3	50,1	51,0	53,6
<i>Markedness</i> [%]	51,0	48,1	58,1	54,7	53,4	57,4

8.2.14.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 58,1 %. A Tabela 8.64 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.64: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	58,8	57,9	66,7	63,0	61,6	65,9
Taxa de Verdadeiras Positivas [%]	52,9	54,1	60,0	60,0	60,7	60,8
Taxa de Verdadeiras Negativas [%]	91,2	90,5	92,9	91,4	90,8	92,6
Valor Preditivo Negativo [%]	91,0	91,2	92,4	92,4	92,4	92,6
<i>F1 Score</i> [%]	57,2	56,8	62,8	61,3	62,2	62,9
<i>Informedness</i> [%]	44,2	44,6	53,0	51,5	51,5	53,4
<i>Markedness</i> [%]	49,9	49,1	59,2	55,4	54,0	58,5

8.2.14.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 52,8 %. A Tabela 8.65 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.65: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	48,7	50,3	63,5	62,2	60,4	68,9
Taxa de Verdadeiras Positivas [%]	72,1	55,4	55,4	53,4	43,5	37,0
Taxa de Verdadeiras Negativas [%]	80,6	86,6	92,6	91,4	94,3	97,7
Valor Preditivo Negativo [%]	93,9	91,1	91,6	91,1	89,7	88,9
<i>F1 Score</i> [%]	58,1	55,6	58,9	55,8	52,8	45,4
<i>Informedness</i> [%]	52,8	42,0	48,1	44,8	37,9	34,7
<i>Markedness</i> [%]	42,7	41,4	55,2	53,4	50,1	57,9

8.2.15 Somente a Obliquidade no Domínio do Tempo

Nesta seção, considera-se apenas a obliquidade do sinal, que foi descrita na Seção 4.5.5, no domínio do tempo.

8.2.15.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.66 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.66: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a obliquidade do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	100,0	100,0	100,0	100,0	100,0	99,9
Taxa de Verdadeiras Positivas [%]	99,9	100,0	100,0	100,0	100,0	100,0
Taxa de Verdadeiras Negativas [%]	100,0	100,0	100,0	100,0	100,0	99,9
Valor Preditivo Negativo [%]	99,9	100,0	100,0	100,0	100,0	100,0
<i>F1 Score</i> [%]	99,9	100,0	100,0	100,0	100,0	99,9
<i>Informedness</i> [%]	99,9	100,0	100,0	100,0	100,0	99,9
<i>Markedness</i> [%]	99,9	100,0	100,0	100,0	100,0	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 85,4 %. A Tabela 8.67 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.67: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,0	87,1	85,7	91,0	90,4	91,2
Taxa de Verdadeiras Positivas [%]	84,8	85,6	82,2	87,2	86,7	85,8
Taxa de Verdadeiras Negativas [%]	95,7	96,6	96,5	97,8	97,7	98,0
Valor Preditivo Negativo [%]	97,1	97,3	96,7	97,6	97,5	97,4
<i>F1 Score</i> [%]	82,9	84,0	82,0	87,0	86,4	85,9
<i>Informedness</i> [%]	80,5	82,3	78,7	85,1	84,4	83,8
<i>Markedness</i> [%]	81,1	84,4	82,4	88,7	87,9	88,6

8.2.15.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 85,8%. A Tabela 8.68 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.68: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	84,1	87,4	86,0	91,2	90,5	90,9
Taxa de Verdadeiras Positivas [%]	85,4	85,7	81,7	87,6	87,6	86,7
Taxa de Verdadeiras Negativas [%]	95,8	96,7	96,6	98,0	97,7	97,9
Valor Preditivo Negativo [%]	97,2	97,3	96,6	97,7	97,7	97,6
<i>F1 Score</i> [%]	83,3	84,3	81,5	87,2	87,0	86,3
<i>Informedness</i> [%]	81,2	82,4	78,4	85,6	85,3	84,6
<i>Markedness</i> [%]	81,4	84,7	82,6	89,0	88,2	88,5

8.2.15.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 79,1%. A Tabela 8.69 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.69: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	68,7	79,8	82,0	89,9	90,0	92,6
Taxa de Verdadeiras Positivas [%]	97,2	86,9	67,3	79,2	73,2	71,0
Taxa de Verdadeiras Negativas [%]	87,2	94,0	97,1	98,1	98,8	99,6
Valor Preditivo Negativo [%]	99,3	97,5	94,1	96,2	95,2	94,9
<i>F1 Score</i> [%]	78,1	81,1	70,0	80,9	77,6	76,2
<i>Informedness</i> [%]	84,4	80,9	64,4	77,3	72,1	70,7
<i>Markedness</i> [%]	68,1	77,3	76,1	86,2	85,2	87,5

8.2.16 Somente a Curtose no Domínio do Tempo

Nesta seção, considera-se apenas a curtose do sinal, que foi descrita na Seção 4.5.6, no domínio do tempo.

8.2.16.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.70 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.70: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a curtose do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,9	100,0	99,9	99,9	100,0	99,9
Taxa de Verdadeiras Positivas [%]	100,0	99,9	100,0	99,9	99,8	100,0
Taxa de Verdadeiras Negativas [%]	99,9	100,0	99,9	100,0	100,0	100,0
Valor Preditivo Negativo [%]	100,0	100,0	100,0	100,0	99,9	100,0
<i>F1 Score</i> [%]	99,9	100,0	99,9	99,9	99,9	99,9
<i>Informedness</i> [%]	99,9	99,9	99,9	99,9	99,8	100,0
<i>Markedness</i> [%]	99,9	100,0	99,9	99,9	99,9	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 70,8 %. A Tabela 8.71 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.71: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	74,0	71,6	79,0	73,5	69,3	81,8
Taxa de Verdadeiras Positivas [%]	73,6	68,4	73,8	67,5	70,8	70,5
Taxa de Verdadeiras Negativas [%]	93,6	93,3	95,2	94,1	92,0	96,5
Valor Preditivo Negativo [%]	95,0	94,0	95,1	93,9	94,3	94,6
<i>F1 Score</i> [%]	72,1	67,5	73,9	68,2	69,7	73,4
<i>Informedness</i> [%]	67,3	61,7	69,0	61,7	62,8	67,1
<i>Markedness</i> [%]	69,1	65,7	74,1	67,5	63,7	76,4

8.2.16.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 71,4 %. A Tabela 8.72 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.72: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	74,4	71,5	79,6	73,9	70,7	81,6
Taxa de Verdadeiras Positivas [%]	73,5	68,6	74,3	69,7	71,9	70,7
Taxa de Verdadeiras Negativas [%]	93,9	93,4	95,3	94,0	92,4	96,6
Valor Preditivo Negativo [%]	95,0	94,0	95,2	94,3	94,5	94,6
<i>F1 Score</i> [%]	72,7	68,2	74,7	69,9	70,2	73,8
<i>Informedness</i> [%]	67,4	62,1	69,7	63,8	64,3	67,3
<i>Markedness</i> [%]	69,4	65,5	74,8	68,3	65,3	76,2

8.2.16.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 60,5 %. A Tabela 8.73 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.73: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio do tempo.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	56,3	56,8	71,7	67,1	66,7	81,0
Taxa de Verdadeiras Positivas [%]	85,1	68,4	60,1	56,4	41,3	51,7
Taxa de Verdadeiras Negativas [%]	82,1	87,0	94,7	93,7	96,1	98,8
Valor Preditivo Negativo [%]	96,8	93,6	92,6	91,9	89,5	91,5
<i>F1 Score</i> [%]	65,1	61,5	64,2	58,8	50,0	59,2
<i>Informedness</i> [%]	67,2	55,4	54,9	50,1	37,5	50,5
<i>Markedness</i> [%]	53,1	50,5	64,4	59,1	56,2	72,5

8.2.17 Somente o Valor RMS no Domínio da Frequência

Nesta seção, considera-se apenas o valor RMS do sinal, que foi descrito na Seção 4.5.1, no domínio da frequência.

8.2.17.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.74 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.74: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o valor RMS do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,8	99,9	99,9	100,0	99,9	99,9
Taxa de Verdadeiras Positivas [%]	99,7	99,9	100,0	99,8	99,9	99,9
Taxa de Verdadeiras Negativas [%]	99,9	99,9	99,9	100,0	99,9	99,9
Valor Preditivo Negativo [%]	99,9	100,0	100,0	99,9	99,9	99,9
<i>F1 Score</i> [%]	99,7	99,9	99,9	99,9	99,9	99,9
<i>Informedness</i> [%]	99,6	99,9	99,9	99,8	99,9	99,9
<i>Markedness</i> [%]	99,8	99,9	99,9	99,9	99,9	99,8

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 72,4%. A Tabela 8.75 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.75: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	66,6	72,8	75,8	80,7	85,5	76,8
Taxa de Verdadeiras Positivas [%]	69,4	69,9	72,0	75,0	80,0	68,3
Taxa de Verdadeiras Negativas [%]	91,0	93,9	94,5	95,6	96,7	95,0
Valor Preditivo Negativo [%]	94,0	94,3	94,8	95,3	96,3	94,1
<i>F1 Score</i> [%]	67,5	70,9	71,3	75,9	80,8	69,3
<i>Informedness</i> [%]	60,4	63,9	66,6	70,7	76,7	63,3
<i>Markedness</i> [%]	60,7	67,2	70,7	76,1	81,8	70,9

8.2.17.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 73,0%. A Tabela 8.76 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.76: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	68,6	74,6	75,0	81,1	85,1	76,6
Taxa de Verdadeiras Positivas [%]	71,3	69,2	72,9	77,0	80,3	67,2
Taxa de Verdadeiras Negativas [%]	91,6	94,6	94,1	95,5	96,4	95,1
Valor Preditivo Negativo [%]	94,4	94,2	94,9	95,7	96,3	93,9
<i>F1 Score</i> [%]	68,9	71,1	71,8	77,1	80,9	68,6
<i>Informedness</i> [%]	62,9	63,8	67,0	72,5	76,8	62,4
<i>Markedness</i> [%]	63,0	68,9	69,9	76,8	81,4	70,6

8.2.17.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 63,3 %. A Tabela 8.77 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.77: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o valor RMS do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	47,9	66,8	66,3	80,9	87,0	76,3
Taxa de Verdadeiras Positivas [%]	89,7	67,1	54,8	64,0	64,2	40,0
Taxa de Verdadeiras Negativas [%]	75,0	92,0	93,5	97,3	98,9	98,9
Valor Preditivo Negativo [%]	97,6	93,7	91,7	93,5	93,6	89,5
<i>F1 Score</i> [%]	60,8	66,6	56,4	69,6	70,7	47,4
<i>Informedness</i> [%]	64,8	59,2	48,4	61,4	63,2	38,9
<i>Markedness</i> [%]	45,5	60,6	58,0	74,5	80,7	65,8

8.2.18 Somente a Média Aritmética no Domínio da Frequência

Nesta seção, considera-se apenas a média aritmética do sinal, que foi descrita na Seção 4.5.2, no domínio da frequência.

8.2.18.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,9 %. A Tabela 8.78 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.78: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a média aritmética do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,8	99,9	100,0	99,9	99,9	99,9
Taxa de Verdadeiras Positivas [%]	99,8	99,9	100,0	99,7	99,9	100,0
Taxa de Verdadeiras Negativas [%]	99,9	99,9	100,0	100,0	99,9	99,9
Valor Preditivo Negativo [%]	99,9	100,0	100,0	99,9	100,0	100,0
<i>F1 Score</i> [%]	99,8	99,9	100,0	99,8	99,9	99,9
<i>Informedness</i> [%]	99,8	99,9	100,0	99,7	99,9	99,9
<i>Markedness</i> [%]	99,7	99,9	100,0	99,9	99,9	99,9

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 74,8 %. A Tabela 8.79 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.79: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	66,7	76,1	76,3	84,7	89,0	79,2
Taxa de Verdadeiras Positivas [%]	68,2	72,2	73,3	78,0	83,0	74,4
Taxa de Verdadeiras Negativas [%]	91,2	94,7	94,6	96,3	97,5	95,3
Valor Preditivo Negativo [%]	93,8	94,8	95,0	95,9	96,8	95,3
<i>F1 Score</i> [%]	66,5	72,8	72,2	78,5	83,7	73,9
<i>Informedness</i> [%]	59,5	66,9	67,9	74,3	80,5	69,8
<i>Markedness</i> [%]	60,5	70,9	71,3	80,7	85,9	74,6

8.2.18.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 75,6%. A Tabela 8.80 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.80: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	68,6	76,7	76,7	85,5	88,8	80,0
Taxa de Verdadeiras Positivas [%]	69,4	73,2	73,6	77,9	85,1	74,5
Taxa de Verdadeiras Negativas [%]	91,7	94,7	94,7	96,7	97,3	95,5
Valor Preditivo Negativo [%]	94,0	94,9	95,1	95,9	97,2	95,3
<i>F1 Score</i> [%]	67,8	73,3	72,9	79,1	85,0	74,2
<i>Informedness</i> [%]	61,1	67,9	68,3	74,6	82,4	70,0
<i>Markedness</i> [%]	62,7	71,7	71,8	81,5	86,1	75,3

8.2.18.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 65,9%. A Tabela 8.81 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.81: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a média aritmética do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	48,5	67,1	69,7	83,3	90,4	80,4
Taxa de Verdadeiras Positivas [%]	90,9	65,6	53,5	65,9	71,4	48,0
Taxa de Verdadeiras Negativas [%]	75,2	92,6	95,0	97,7	99,1	99,2
Valor Preditivo Negativo [%]	97,8	93,5	91,6	93,9	94,9	90,9
<i>F1 Score</i> [%]	61,4	66,5	57,0	70,7	76,6	54,2
<i>Informedness</i> [%]	66,1	58,3	48,6	63,6	70,6	47,2
<i>Markedness</i> [%]	46,4	60,6	61,3	77,2	85,3	71,4

8.2.19 Somente a Variância no Domínio da Frequência

Nesta seção, considera-se apenas a variância do sinal, que foi descrita na Seção 4.5.3, no domínio da frequência.

8.2.19.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,7 %. A Tabela 8.82 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.82: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas a variância do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,5	99,8	99,6	99,9	99,9	99,7
Taxa de Verdadeiras Positivas [%]	99,5	99,9	99,9	99,4	99,9	99,7
Taxa de Verdadeiras Negativas [%]	99,8	99,9	99,9	100,0	99,9	99,9
Valor Preditivo Negativo [%]	99,9	99,9	99,9	99,8	99,9	99,9
<i>F1 Score</i> [%]	99,5	99,8	99,7	99,6	99,9	99,7
<i>Informedness</i> [%]	99,4	99,8	99,8	99,4	99,8	99,6
<i>Markedness</i> [%]	99,5	99,7	99,6	99,8	99,9	99,7

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 69,5 %. A Tabela 8.83 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.83: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	64,1	70,5	72,9	77,2	83,3	74,1
Taxa de Verdadeiras Positivas [%]	67,2	66,6	71,4	72,0	75,1	65,0
Taxa de Verdadeiras Negativas [%]	90,7	93,4	93,4	94,8	96,4	94,6
Valor Preditivo Negativo [%]	93,6	93,7	94,6	94,7	95,4	93,5
<i>F1 Score</i> [%]	65,3	68,2	69,8	73,3	77,0	67,0
<i>Informedness</i> [%]	57,9	60,0	64,8	66,8	71,6	59,6
<i>Markedness</i> [%]	57,8	64,2	67,6	72,0	78,8	67,6

8.2.19.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 69,8%. A Tabela 8.84 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.84: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	65,3	72,9	71,9	77,1	82,0	74,0
Taxa de Verdadeiras Positivas [%]	69,5	66,8	70,3	72,4	75,0	64,7
Taxa de Verdadeiras Negativas [%]	90,7	94,2	93,4	94,7	96,1	94,5
Valor Preditivo Negativo [%]	94,0	93,8	94,4	94,8	95,3	93,4
<i>F1 Score</i> [%]	66,9	68,8	69,4	73,7	76,5	66,8
<i>Informedness</i> [%]	60,2	61,0	63,8	67,2	71,1	59,2
<i>Markedness</i> [%]	59,4	66,7	66,4	71,9	77,4	67,5

8.2.19.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 57,7%. A Tabela 8.85 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.85: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a variância do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	44,0	61,9	58,9	74,1	82,6	72,0
Taxa de Verdadeiras Positivas [%]	88,4	64,4	48,7	54,2	56,1	34,2
Taxa de Verdadeiras Negativas [%]	72,2	90,5	92,2	96,6	98,6	98,8
Valor Preditivo Negativo [%]	97,1	93,1	90,5	91,7	92,2	88,6
<i>F1 Score</i> [%]	57,7	63,3	52,1	61,2	63,6	42,4
<i>Informedness</i> [%]	60,7	54,9	41,0	50,8	54,7	33,0
<i>Markedness</i> [%]	41,2	55,1	49,5	65,9	74,8	60,7

8.2.20 Somente o Desvio Padrão no Domínio da Frequência

Nesta seção, considera-se apenas o desvio padrão do sinal, que foi descrito na Seção 4.5.4, no domínio da frequência.

8.2.20.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 99,8%. A Tabela 8.86 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de treinamento.

Tabela 8.86: Desempenho do modelo de regressão logística contra o conjunto de treinamento quando se considera apenas o desvio padrão do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	99,8	99,9	99,8	99,9	99,9	99,8
Taxa de Verdadeiras Positivas [%]	99,6	99,9	100,0	99,7	99,9	99,9
Taxa de Verdadeiras Negativas [%]	99,9	99,9	99,9	100,0	99,9	99,9
Valor Preditivo Negativo [%]	99,9	100,0	100,0	99,9	100,0	99,9
<i>F1 Score</i> [%]	99,7	99,9	99,9	99,8	99,9	99,9
<i>Informedness</i> [%]	99,6	99,9	99,9	99,7	99,9	99,9
<i>Markedness</i> [%]	99,8	99,9	99,8	99,9	99,9	99,8

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 72,3 %. A Tabela 8.87 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.87: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	66,9	74,2	76,1	79,3	83,7	77,4
Taxa de Verdadeiras Positivas [%]	69,6	69,5	71,8	75,5	79,3	68,3
Taxa de Verdadeiras Negativas [%]	91,1	94,3	94,6	95,1	96,3	95,2
Valor Preditivo Negativo [%]	94,1	94,2	94,7	95,4	96,1	94,1
<i>F1 Score</i> [%]	67,4	71,2	71,3	75,4	79,7	69,9
<i>Informedness</i> [%]	60,8	63,9	66,4	70,7	75,6	63,6
<i>Markedness</i> [%]	61,0	68,4	70,9	74,8	79,8	71,6

8.2.20.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 73,1 %. A Tabela 8.88 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.88: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	68,3	75,2	75,1	80,7	84,1	77,9
Taxa de Verdadeiras Positivas [%]	71,7	70,1	73,0	75,6	80,0	68,3
Taxa de Verdadeiras Negativas [%]	91,6	94,6	94,1	95,5	96,3	95,4
Valor Preditivo Negativo [%]	94,5	94,4	94,9	95,4	96,2	94,1
<i>F1 Score</i> [%]	69,6	71,4	71,7	76,4	80,4	70,1
<i>Informedness</i> [%]	63,3	64,8	67,1	71,2	76,3	63,8
<i>Markedness</i> [%]	62,8	69,6	70,1	76,1	80,4	72,1

8.2.20.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 63,3%. A Tabela 8.89 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.89: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o desvio padrão do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	48,4	68,0	65,7	79,3	85,3	76,6
Taxa de Verdadeiras Positivas [%]	89,8	68,2	56,0	62,9	62,4	40,6
Taxa de Verdadeiras Negativas [%]	75,6	92,2	93,4	96,9	98,7	99,0
Valor Preditivo Negativo [%]	97,7	94,0	91,9	93,3	93,3	89,6
<i>F1 Score</i> [%]	61,4	67,4	57,6	68,2	69,2	48,4
<i>Informedness</i> [%]	65,5	60,5	49,5	59,9	61,1	39,6
<i>Markedness</i> [%]	46,1	62,0	57,7	72,6	78,6	66,3

8.2.21 Somente a Obliquidade no Domínio da Frequência

Nesta seção, considera-se apenas a obliquidade do sinal, que foi descrita na Seção 4.5.5, no domínio da frequência.

8.2.21.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 92,7%. A Tabela 8.90 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.90: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	91,2	94,7	92,0	96,0	96,1	96,5
Taxa de Verdadeiras Positivas [%]	90,3	91,0	88,8	94,0	95,6	96,4
Taxa de Verdadeiras Negativas [%]	97,6	98,6	97,9	99,0	98,9	99,0
Valor Preditivo Negativo [%]	98,2	98,3	97,9	98,9	99,2	99,3
<i>F1 Score</i> [%]	88,8	91,2	88,2	93,7	94,8	95,7
<i>Informedness</i> [%]	88,0	89,7	86,7	93,0	94,5	95,4
<i>Markedness</i> [%]	89,4	93,1	89,9	94,9	95,3	95,9

8.2.21.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 93,1 %. A Tabela 8.91 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.91: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,7	95,4	91,5	96,4	96,1	97,0
Taxa de Verdadeiras Positivas [%]	92,0	91,4	89,1	93,7	96,1	96,3
Taxa de Verdadeiras Negativas [%]	97,9	98,8	97,8	99,0	98,9	99,1
Valor Preditivo Negativo [%]	98,5	98,4	98,0	98,8	99,3	99,3
<i>F1 Score</i> [%]	90,6	91,8	88,3	93,7	95,2	95,8
<i>Informedness</i> [%]	90,0	90,2	87,0	92,8	95,0	95,4
<i>Markedness</i> [%]	91,2	93,8	89,5	95,3	95,4	96,3

8.2.21.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 87,3 %. A Tabela 8.92 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.92: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a obliquidade do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	79,0	87,3	88,4	94,8	96,7	97,7
Taxa de Verdadeiras Positivas [%]	98,7	92,1	73,9	85,2	86,6	87,2
Taxa de Verdadeiras Negativas [%]	92,0	95,9	98,3	99,1	99,6	99,7
Valor Preditivo Negativo [%]	99,7	98,5	95,3	97,3	97,6	97,7
<i>F1 Score</i> [%]	85,7	87,2	77,2	87,2	89,2	90,1
<i>Informedness</i> [%]	90,7	88,1	72,2	84,3	86,2	86,9
<i>Markedness</i> [%]	78,8	85,8	83,7	92,2	94,3	95,5

8.2.22 Somente a Curtose no Domínio da Frequência

Nesta seção, considera-se apenas a curtose do sinal, que foi descrita na Seção 4.5.6, no domínio da frequência.

8.2.22.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 88,1 %. A Tabela 8.93 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.93: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	85,0	91,1	88,6	92,9	91,9	93,8
Taxa de Verdadeiras Positivas [%]	85,2	88,9	84,7	88,4	90,7	90,5
Taxa de Verdadeiras Negativas [%]	96,0	97,7	97,3	98,2	97,9	98,4
Valor Preditivo Negativo [%]	97,2	97,9	97,2	97,8	98,3	98,2
<i>F1 Score</i> [%]	82,8	88,0	84,2	88,6	89,4	90,5
<i>Informedness</i> [%]	81,2	86,6	82,1	86,6	88,6	89,0
<i>Markedness</i> [%]	82,3	89,1	85,8	90,8	90,2	92,0

8.2.22.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 88,4 %. A Tabela 8.94 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.94: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	86,7	91,6	88,3	93,3	92,1	93,3
Taxa de Verdadeiras Positivas [%]	85,7	88,9	85,5	88,0	92,0	90,2
Taxa de Verdadeiras Negativas [%]	96,3	97,8	97,2	98,4	97,9	98,3
Valor Preditivo Negativo [%]	97,3	97,9	97,3	97,8	98,5	98,2
<i>F1 Score</i> [%]	83,9	88,4	84,7	88,5	90,3	90,3
<i>Informedness</i> [%]	82,1	86,8	82,7	86,4	89,9	88,5
<i>Markedness</i> [%]	84,0	89,6	85,7	91,1	90,7	91,5

8.2.22.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 78,9 %. A Tabela 8.95 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.95: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas a curtose do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	69,3	80,7	80,4	86,8	92,0	92,9
Taxa de Verdadeiras Positivas [%]	95,1	86,5	64,9	78,2	75,0	73,5
Taxa de Verdadeiras Negativas [%]	87,5	94,1	96,9	97,3	99,1	99,5
Valor Preditivo Negativo [%]	99,0	97,5	93,7	96,0	95,5	95,3
<i>F1 Score</i> [%]	77,5	80,3	67,9	79,3	79,6	78,3
<i>Informedness</i> [%]	82,7	80,6	61,9	75,5	74,1	73,0
<i>Markedness</i> [%]	68,3	78,2	74,1	82,8	87,6	88,2

8.2.23 Somente o Centroe Espectral no Domínio da Frequência

Nesta seção, considera-se apenas o centroe espectral do sinal, que foi descrito na Seção 4.5.7, no domínio da frequência.

8.2.23.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,0%. A Tabela 8.96 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.96: Desempenho do modelo de regressão logística contra o conjunto de teste quando se considera apenas o centroe espectral do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,2	97,1	95,3	97,9	99,4	98,4
Taxa de Verdadeiras Positivas [%]	95,6	92,9	93,4	97,3	99,0	97,6
Taxa de Verdadeiras Negativas [%]	98,2	99,2	98,7	99,5	99,8	99,5
Valor Preditivo Negativo [%]	99,2	98,7	98,8	99,5	99,8	99,5
<i>F1 Score</i> [%]	93,7	93,7	93,0	97,1	99,0	97,5
<i>Informedness</i> [%]	93,9	92,1	92,2	96,8	98,8	97,2
<i>Markedness</i> [%]	93,4	95,8	94,1	97,5	99,2	98,0

8.2.23.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,3%. A Tabela 8.97 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.97: Desempenho do modelo de rede neural contra o conjunto de teste quando se considera apenas o centroide espectral do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,6	97,2	95,6	98,2	99,7	98,4
Taxa de Verdadeiras Positivas [%]	96,2	93,4	93,3	97,5	99,2	98,0
Taxa de Verdadeiras Negativas [%]	98,4	99,2	98,8	99,5	99,9	99,5
Valor Preditivo Negativo [%]	99,3	98,7	98,7	99,5	99,8	99,6
<i>F1 Score</i> [%]	94,3	94,1	93,1	97,3	99,3	97,8
<i>Informedness</i> [%]	94,6	92,6	92,1	97,1	99,1	97,6
<i>Markedness</i> [%]	93,9	96,0	94,4	97,8	99,6	98,1

8.2.23.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 93,9%. A Tabela 8.98 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

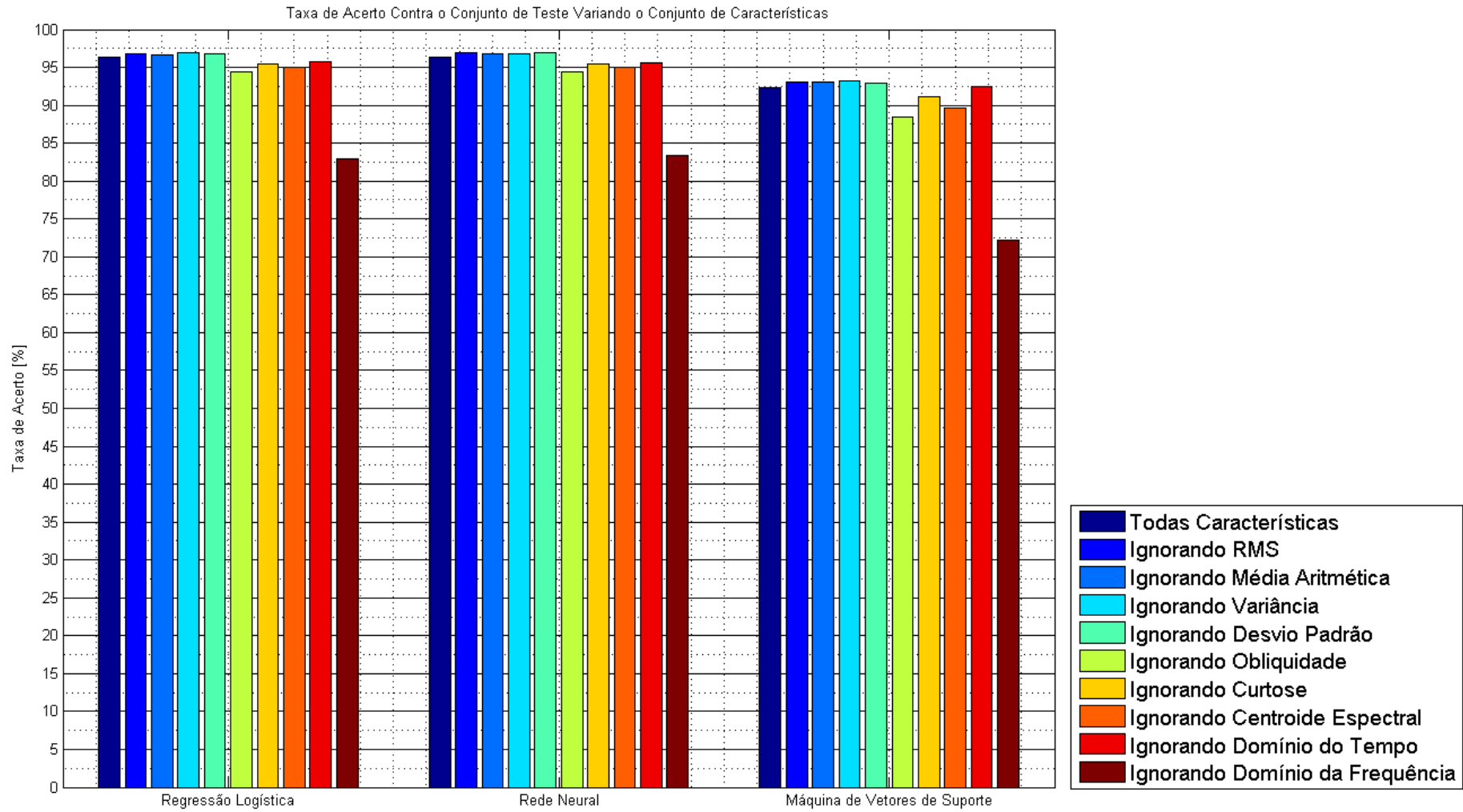
Tabela 8.98: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando se considera apenas o centroide espectral do sinal no domínio da frequência.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	87,2	92,6	96,3	98,2	99,9	99,4
Taxa de Verdadeiras Positivas [%]	99,5	94,9	81,5	94,3	98,0	95,1
Taxa de Verdadeiras Negativas [%]	95,6	97,7	99,7	99,6	99,9	99,9
Valor Preditivo Negativo [%]	99,9	99,0	96,7	98,9	99,6	99,1
<i>F1 Score</i> [%]	91,7	92,3	85,5	95,1	98,7	96,3
<i>Informedness</i> [%]	95,2	92,6	81,3	94,0	98,0	95,0
<i>Markedness</i> [%]	87,2	91,6	93,0	97,2	99,6	98,5

8.2.24 Análise de Desempenho Variando o Conjunto de Características

A Figura 8.3 mostra a taxa de acerto de cada um dos modelos para conjuntos nos quais se ignora uma das características descritas na Seção 4.5 ou todas as características de um determinado domínio (tempo ou frequência). Fica claro, através dessa figura, que nenhuma das características consideradas é, na presença de todas as outras, insubstituível. No entanto, é visível a importância das características pertencentes ao domínio da frequência para a classificação dos movimentos desta base de dados. A taxa de acerto média do sistema é pouco afetada quando se ignora as características pertencentes ao domínio do tempo, porém, ao se ignorar as características do domínio da frequência, a variação na taxa de acerto média dos modelos é evidente.

Figura 8.3: Taxa de acerto variando o conjunto de características (ignorando características).



A Figura 8.4 examina, separadamente, a taxa de acerto dos modelos para conjuntos contendo apenas uma das características (em apenas um dos domínios) descritas na Seção 4.5. Este gráfico mostra que, enquanto algumas características (como a média aritmética do sinal no domínio do tempo) não conseguem, sozinhas, classificar, eficientemente, os movimentos considerados da base de dados, outras (como a obliquidade e o centroide espectral no domínio da frequência) conseguem, sozinhas, boas taxas de acerto médias. A Figura 8.5 - que mostra a métrica *informedness* para cada classe de movimento, para cada configuração e para cada modelo de aprendizado - reforça este fato, mostrando valores muito baixos (ou até mesmo negativos) para o caso que considera somente a média aritmética do sinal no domínio do tempo e valores elevados para os casos que consideram somente a obliquidade ou o centroide espectral no domínio da frequência.

Como visto na Seção 8.1.7, percebe-se, observando as Figuras 8.3 e 8.4, que, também para o conjunto de testes considerados nesta seção, os modelos de regressão logística e rede neural possuem taxas de acerto bastante semelhantes, e o modelo de máquina de vetores de suporte apresenta uma taxa de acerto inferior às taxas dos outros modelos em cada um dos casos considerados.

Figura 8.4: Taxa de acerto variando o conjunto de características (utilizando somente uma das características).

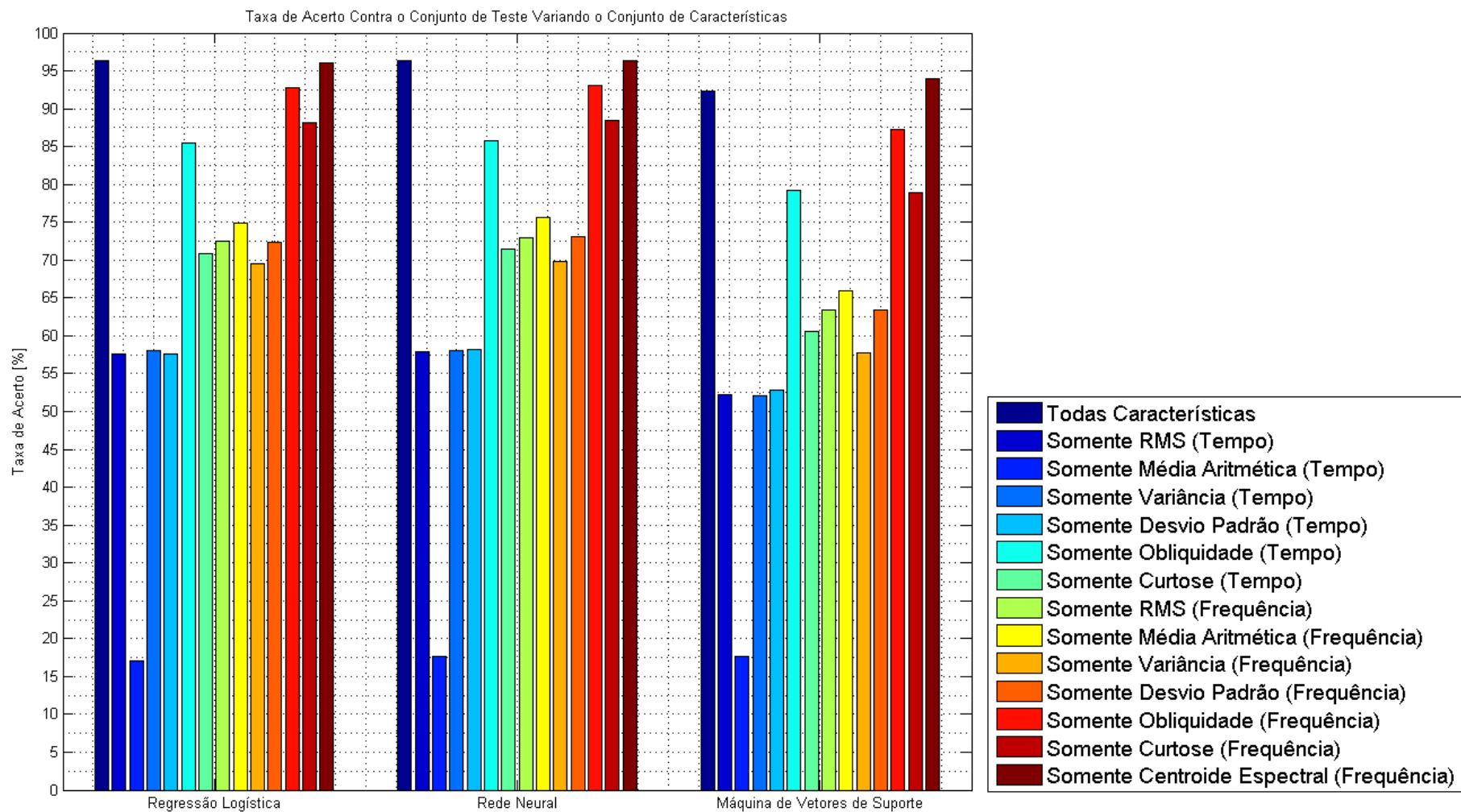
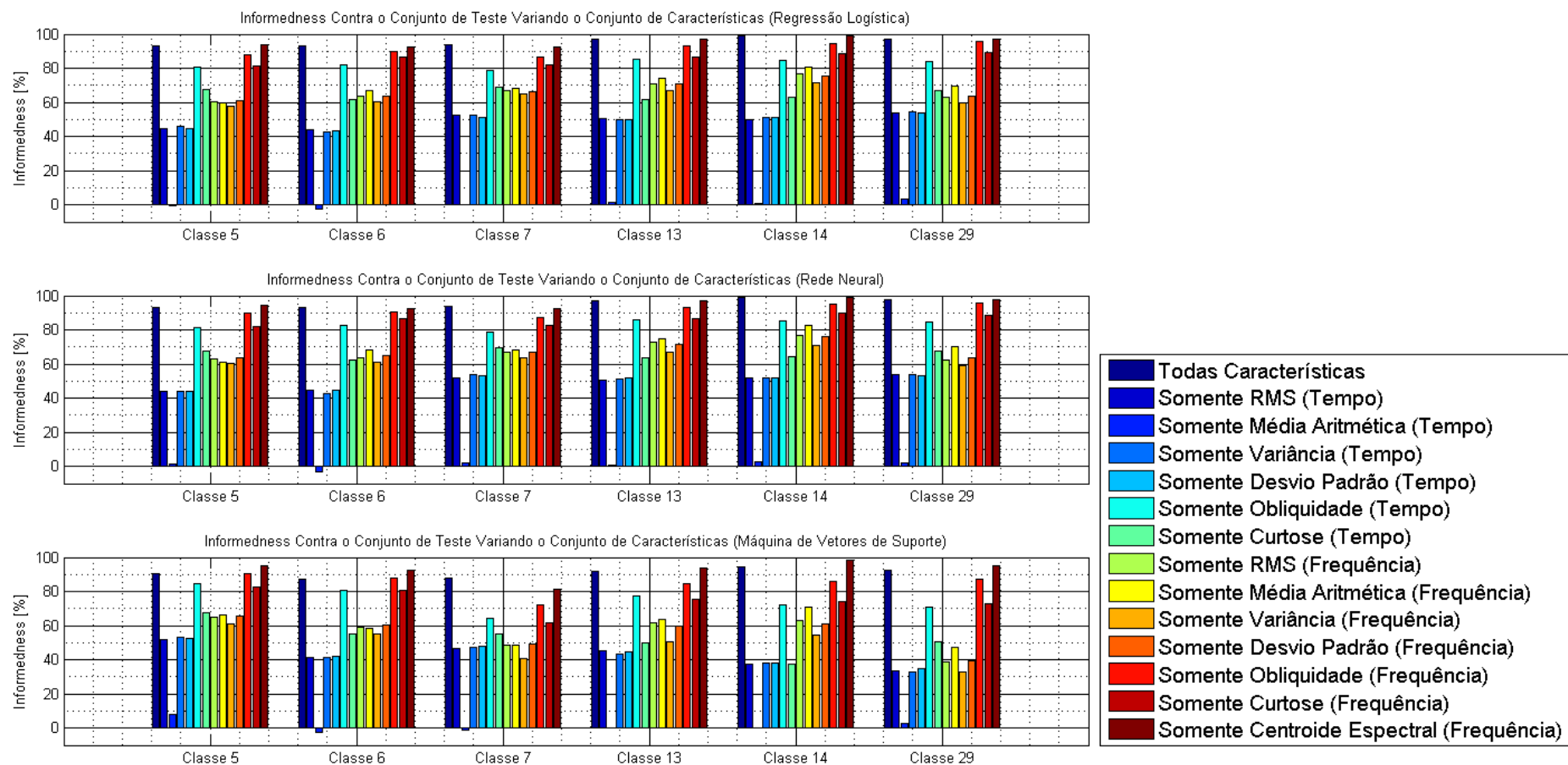


Figura 8.5: *Informedness* variando o conjunto de características (utilizando somente uma das características).



8.3 Desempenho Variando o Tamanho do Conjunto de Treinamento

Nesta seção, avalia-se o desempenho dos modelos de aprendizado de máquina para tamanhos variáveis de conjunto de treinamento. Os seguintes conjuntos são avaliados:

- conjunto de treinamento com 90 % dos exemplos disponíveis (Seção 8.3.1);
- conjunto de treinamento com 80 % dos exemplos disponíveis (Seção 8.3.1);
- conjunto de treinamento com 50 % dos exemplos disponíveis (Seção 8.3.1);
- conjunto de treinamento com 40 % dos exemplos disponíveis (Seção 8.3.1);
- conjunto de treinamento com 20 % dos exemplos disponíveis (Seção 8.3.1);

Após a exibição de resultados, na Seção 8.3.6, faz-se uma avaliação de desempenho geral, comparando todos os conjuntos considerados.

8.3.1 Utilizando 90% dos Exemplos para o Treinamento

Nesta seção, o treinamento dos modelos é realizado com 90 % dos exemplos; e o teste, com os outros 10 % dos exemplos.

8.3.1.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,7 %. A Tabela 8.99 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.99: Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,8	95,8	94,6	98,2	99,4	98,3
Taxa de Verdadeiras Positivas [%]	95,2	93,9	95,6	98,0	99,3	98,1
Taxa de Verdadeiras Negativas [%]	98,4	99,5	98,6	99,7	99,9	99,7
Valor Preditivo Negativo [%]	99,2	98,9	99,2	99,6	99,8	99,6
<i>F1 Score</i> [%]	93,1	93,5	93,7	97,6	99,1	97,6
<i>Informedness</i> [%]	93,7	93,4	94,3	97,7	99,2	97,8
<i>Markedness</i> [%]	93,0	94,8	93,8	97,9	99,3	98,0

8.3.1.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 97,0 %. A Tabela 8.100 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.100: Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,7	96,7	94,9	98,3	99,4	98,2
Taxa de Verdadeiras Positivas [%]	95,4	95,2	96,0	97,9	99,4	98,0
Taxa de Verdadeiras Negativas [%]	98,7	99,6	98,7	99,7	99,9	99,6
Valor Preditivo Negativo [%]	99,2	99,2	99,3	99,6	99,9	99,6
<i>F1 Score</i> [%]	93,7	94,8	94,2	97,6	99,3	97,5
<i>Informedness</i> [%]	94,2	94,9	94,7	97,6	99,3	97,7
<i>Markedness</i> [%]	93,9	95,9	94,2	97,9	99,4	97,9

8.3.1.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 93,3%. A Tabela 8.101 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.101: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 90% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	89,1	90,9	91,4	94,9	97,2	97,0
Taxa de Verdadeiras Positivas [%]	95,7	89,4	91,1	93,2	95,9	94,4
Taxa de Verdadeiras Negativas [%]	96,1	98,4	98,3	99,3	99,7	99,9
Valor Preditivo Negativo [%]	99,2	98,2	98,5	98,8	99,3	99,0
<i>F1 Score</i> [%]	90,3	87,8	88,8	92,4	95,6	94,2
<i>Informedness</i> [%]	91,9	87,9	89,4	92,5	95,6	94,3
<i>Markedness</i> [%]	88,4	89,1	89,9	93,8	96,6	96,0

8.3.2 Utilizando 80% dos Exemplos para o Treinamento

Nesta seção, o treinamento dos modelos é realizado com 80% dos exemplos; e o teste, com os outros 20% dos exemplos.

8.3.2.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 96,3%. A Tabela 8.102 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.102: Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,3	97,8	95,1	98,6	99,3	98,7
Taxa de Verdadeiras Positivas [%]	94,8	93,9	95,0	97,5	99,1	97,3
Taxa de Verdadeiras Negativas [%]	98,3	99,4	98,6	99,6	99,8	99,6
Valor Preditivo Negativo [%]	99,0	98,8	99,0	99,5	99,8	99,5
<i>F1 Score</i> [%]	93,4	94,7	93,8	97,5	99,0	97,5
<i>Informedness</i> [%]	93,1	93,3	93,6	97,1	98,9	97,0
<i>Markedness</i> [%]	93,4	96,7	94,2	98,2	99,1	98,3

8.3.2.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 96,3%. A Tabela 8.103 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.103: Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	94,8	97,6	94,8	98,7	99,1	98,9
Taxa de Verdadeiras Positivas [%]	94,6	93,6	95,0	97,4	99,2	97,9
Taxa de Verdadeiras Negativas [%]	98,5	99,4	98,5	99,6	99,7	99,7
Valor Preditivo Negativo [%]	99,0	98,8	99,1	99,5	99,8	99,6
<i>F1 Score</i> [%]	93,5	94,6	93,7	97,5	98,9	97,9
<i>Informedness</i> [%]	93,1	93,1	93,5	97,0	98,9	97,6
<i>Markedness</i> [%]	93,8	96,4	93,8	98,2	98,9	98,5

8.3.2.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 92,2%. A Tabela 8.104 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.104: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 80% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	86,8	92,5	92,5	96,7	98,1	98,1
Taxa de Verdadeiras Positivas [%]	94,9	88,9	89,5	92,4	94,9	92,9
Taxa de Verdadeiras Negativas [%]	95,5	98,3	98,1	99,2	99,6	99,7
Valor Preditivo Negativo [%]	99,0	97,9	98,1	98,6	99,0	98,7
<i>F1 Score</i> [%]	88,8	88,9	88,8	93,1	95,5	94,1
<i>Informedness</i> [%]	90,4	87,2	87,7	91,7	94,4	92,6
<i>Markedness</i> [%]	85,9	90,5	90,6	95,3	97,2	96,9

8.3.3 Utilizando 50% dos Exemplos para o Treinamento

Nesta seção, o treinamento dos modelos é realizado com 50 % dos exemplos; e o teste, com os outros 50 % dos exemplos.

8.3.3.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 95,1 %. A Tabela 8.105 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.105: Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	92,4	96,5	93,6	97,8	98,6	98,2
Taxa de Verdadeiras Positivas [%]	93,4	92,4	93,5	96,2	98,0	96,8
Taxa de Verdadeiras Negativas [%]	97,9	99,1	98,3	99,4	99,6	99,5
Valor Preditivo Negativo [%]	98,7	98,5	98,8	99,3	99,6	99,4
<i>F1 Score</i> [%]	91,8	93,5	92,4	96,5	98,0	97,0
<i>Informedness</i> [%]	91,4	91,6	91,8	95,7	97,6	96,4
<i>Markedness</i> [%]	91,2	95,1	92,4	97,1	98,2	97,6

8.3.3.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 95,3 %. A Tabela 8.106 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.106: Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	93,3	96,4	93,9	97,7	98,7	98,2
Taxa de Verdadeiras Positivas [%]	93,3	93,0	93,8	96,4	98,6	96,9
Taxa de Verdadeiras Negativas [%]	98,2	99,1	98,4	99,4	99,6	99,5
Valor Preditivo Negativo [%]	98,7	98,7	98,8	99,3	99,7	99,4
<i>F1 Score</i> [%]	92,3	93,8	92,8	96,5	98,4	97,0
<i>Informedness</i> [%]	91,5	92,1	92,2	95,8	98,2	96,4
<i>Markedness</i> [%]	92,1	95,1	92,8	97,1	98,4	97,6

8.3.3.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 90,5%. A Tabela 8.107 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.107: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 50% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	83,4	91,3	91,9	96,0	97,4	97,7
Taxa de Verdadeiras Positivas [%]	94,0	87,4	87,2	91,0	93,4	90,2
Taxa de Verdadeiras Negativas [%]	94,6	97,9	97,9	99,1	99,3	99,6
Valor Preditivo Negativo [%]	98,8	97,6	97,6	98,3	98,7	98,2
<i>F1 Score</i> [%]	86,7	87,5	87,3	92,2	94,4	92,1
<i>Informedness</i> [%]	88,7	85,3	85,2	90,1	92,8	89,8
<i>Markedness</i> [%]	82,2	88,9	89,5	94,4	96,2	96,0

8.3.4 Utilizando 40% dos Exemplos para o Treinamento

Nesta seção, o treinamento dos modelos é realizado com 40% dos exemplos; e o teste, com os outros 60% dos exemplos.

8.3.4.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 92,3%. A Tabela 8.108 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.108: Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	88,4	93,3	91,0	96,4	96,8	97,2
Taxa de Verdadeiras Positivas [%]	90,5	89,4	90,4	93,2	96,3	93,9
Taxa de Verdadeiras Negativas [%]	96,9	98,4	97,7	99,1	99,2	99,3
Valor Preditivo Negativo [%]	98,1	98,0	98,2	98,7	99,3	98,8
<i>F1 Score</i> [%]	88,0	90,2	89,2	93,9	96,1	94,6
<i>Informedness</i> [%]	87,4	87,8	88,1	92,4	95,5	93,2
<i>Markedness</i> [%]	86,6	91,3	89,2	95,1	96,1	96,1

8.3.4.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 92,8%. A Tabela 8.109 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.109: Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	89,6	94,0	91,2	96,4	96,9	97,1
Taxa de Verdadeiras Positivas [%]	89,9	90,3	91,2	93,7	97,0	94,8
Taxa de Verdadeiras Negativas [%]	97,3	98,6	97,7	99,1	99,2	99,2
Valor Preditivo Negativo [%]	98,0	98,1	98,3	98,8	99,4	99,0
<i>F1 Score</i> [%]	88,4	91,1	89,9	94,3	96,5	95,2
<i>Informedness</i> [%]	87,2	88,9	89,0	92,8	96,2	94,1
<i>Markedness</i> [%]	87,7	92,2	89,5	95,2	96,3	96,1

8.3.4.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 86,9%. A Tabela 8.110 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.110: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 40% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	78,1	89,0	89,0	94,0	95,7	95,6
Taxa de Verdadeiras Positivas [%]	90,9	83,7	84,6	86,9	90,0	84,9
Taxa de Verdadeiras Negativas [%]	92,9	97,4	97,2	98,6	98,9	99,1
Valor Preditivo Negativo [%]	98,2	96,9	97,1	97,6	98,1	97,2
<i>F1 Score</i> [%]	81,9	83,8	84,1	88,4	91,2	87,4
<i>Informedness</i> [%]	83,9	81,1	81,8	85,6	88,9	84,1
<i>Markedness</i> [%]	76,3	86,0	86,1	91,6	93,8	92,8

8.3.5 Utilizando 20% dos Exemplos para o Treinamento

Nesta seção, o treinamento dos modelos é realizado com 20 % dos exemplos; e o teste, com os outros 80 % dos exemplos.

8.3.5.1 Regressão Logística

Contra o conjunto de treinamento, o modelo de regressão logística apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de regressão logística apresentou uma taxa de acerto de 82,2 %. A Tabela 8.111 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.111: Desempenho do modelo de regressão logística contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	76,1	84,2	81,6	90,0	90,5	91,7
Taxa de Verdadeiras Positivas [%]	79,0	78,3	80,2	83,9	88,9	83,3
Taxa de Verdadeiras Negativas [%]	93,4	96,3	95,4	97,7	97,7	98,0
Valor Preditivo Negativo [%]	96,0	95,9	96,3	97,0	97,9	96,9
<i>F1 Score</i> [%]	74,2	78,1	77,8	84,5	87,7	84,4
<i>Informedness</i> [%]	72,4	74,7	75,7	81,6	86,6	81,3
<i>Markedness</i> [%]	72,1	80,2	77,9	87,0	88,5	88,6

8.3.5.2 Redes Neurais

Contra o conjunto de treinamento, o modelo de rede neural apresentou uma taxa de acerto de 100,0 %. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de rede neural apresentou uma taxa de acerto de 85,1 %. A Tabela 8.112 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.112: Desempenho do modelo de rede neural contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.

Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	79,8	86,7	83,2	91,0	91,3	92,1
Taxa de Verdadeiras Positivas [%]	79,8	81,0	83,2	86,4	92,2	87,7
Taxa de Verdadeiras Negativas [%]	95,0	97,0	96,0	98,0	97,8	98,1
Valor Preditivo Negativo [%]	96,1	96,4	96,8	97,4	98,5	97,7
<i>F1 Score</i> [%]	77,3	81,6	81,0	87,0	90,5	88,0
<i>Informedness</i> [%]	74,9	78,1	79,2	84,4	90,0	85,8
<i>Markedness</i> [%]	76,0	83,1	80,0	88,5	89,8	89,8

8.3.5.3 Máquina de Vetores de Suporte

Contra o conjunto de treinamento, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 100,0%. A tabela para este caso não é exibida, pois não contém nenhuma informação adicional.

Contra o conjunto de teste, o modelo de máquina de vetores de suporte apresentou uma taxa de acerto de 76,5%. A Tabela 8.113 mostra, mais detalhadamente, o desempenho do modelo contra o conjunto de teste.

Tabela 8.113: Desempenho do modelo de máquina de vetores de suporte contra o conjunto de teste quando treinando o modelo com 20% dos exemplos.

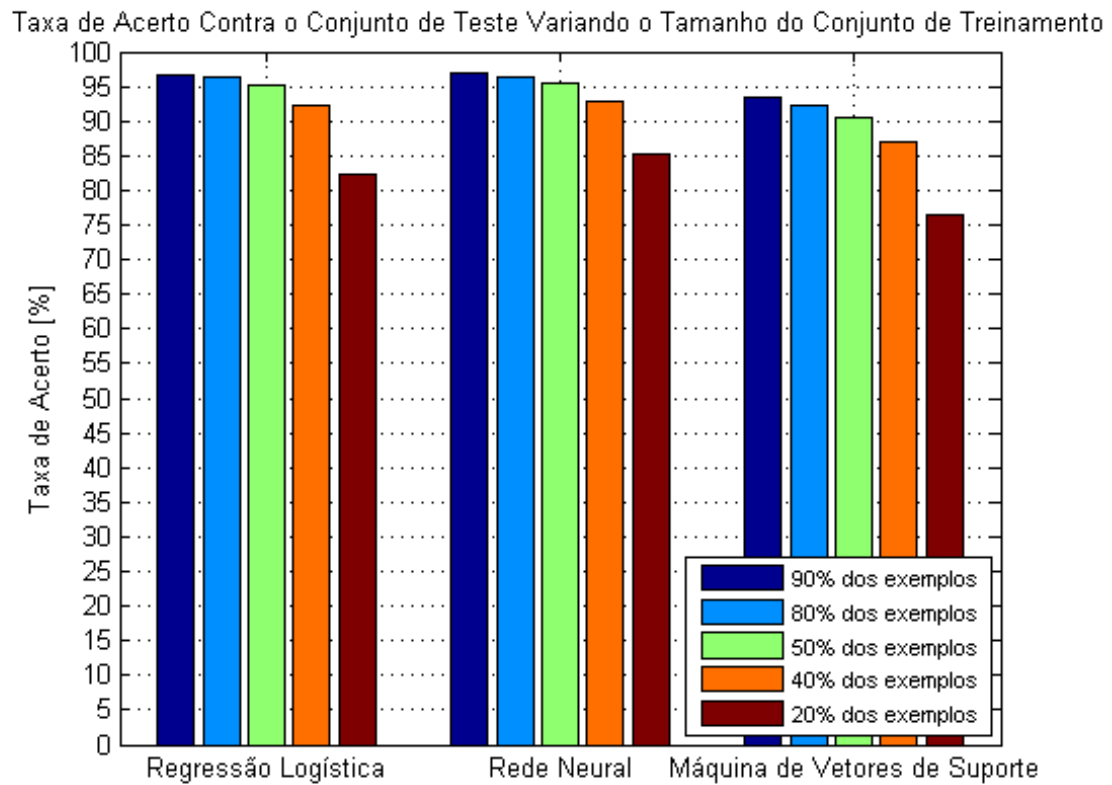
Classe	5	6	7	13	14	29
Valor Preditivo Positivo [%]	66,6	80,9	80,5	87,0	90,3	90,0
Taxa de Verdadeiras Positivas [%]	82,0	72,0	75,5	76,9	80,7	72,3
Taxa de Verdadeiras Negativas [%]	88,5	95,5	94,9	97,1	97,7	98,0
Valor Preditivo Negativo [%]	96,5	94,8	95,5	95,8	96,5	95,0
<i>F1 Score</i> [%]	69,2	71,1	73,2	77,4	81,5	74,9
<i>Informedness</i> [%]	70,5	67,5	70,4	74,0	78,4	70,3
<i>Markedness</i> [%]	63,1	75,8	76,0	82,9	86,8	85,1

8.3.6 Análise de Desempenho Variando o Tamanho do Conjunto de Treinamento

A Figura 8.6 compara, graficamente, para cada um dos modelos de aprendizado considerados, a taxa de acerto contra o conjunto de teste dos algoritmos para tamanhos diferentes de conjunto de treinamento. Como se deveria esperar, a taxa de acerto média de todos os algoritmos diminui com a redução do tamanho do conjunto de treinamento. Novamente, os modelos de regressão logística e de rede neural apresentam taxas de acerto muito próximas - exceto para o último caso (onde se utiliza 20% do conjunto de exemplos para o treinamento e os outros 80% para o teste), no qual o modelo de regressão logística apresenta uma taxa de acerto claramente menor. O modelo de máquina de vetores de suporte, por outro lado, apresenta taxas de acerto visivelmente menores que os outros modelos para cada um dos casos considerados.

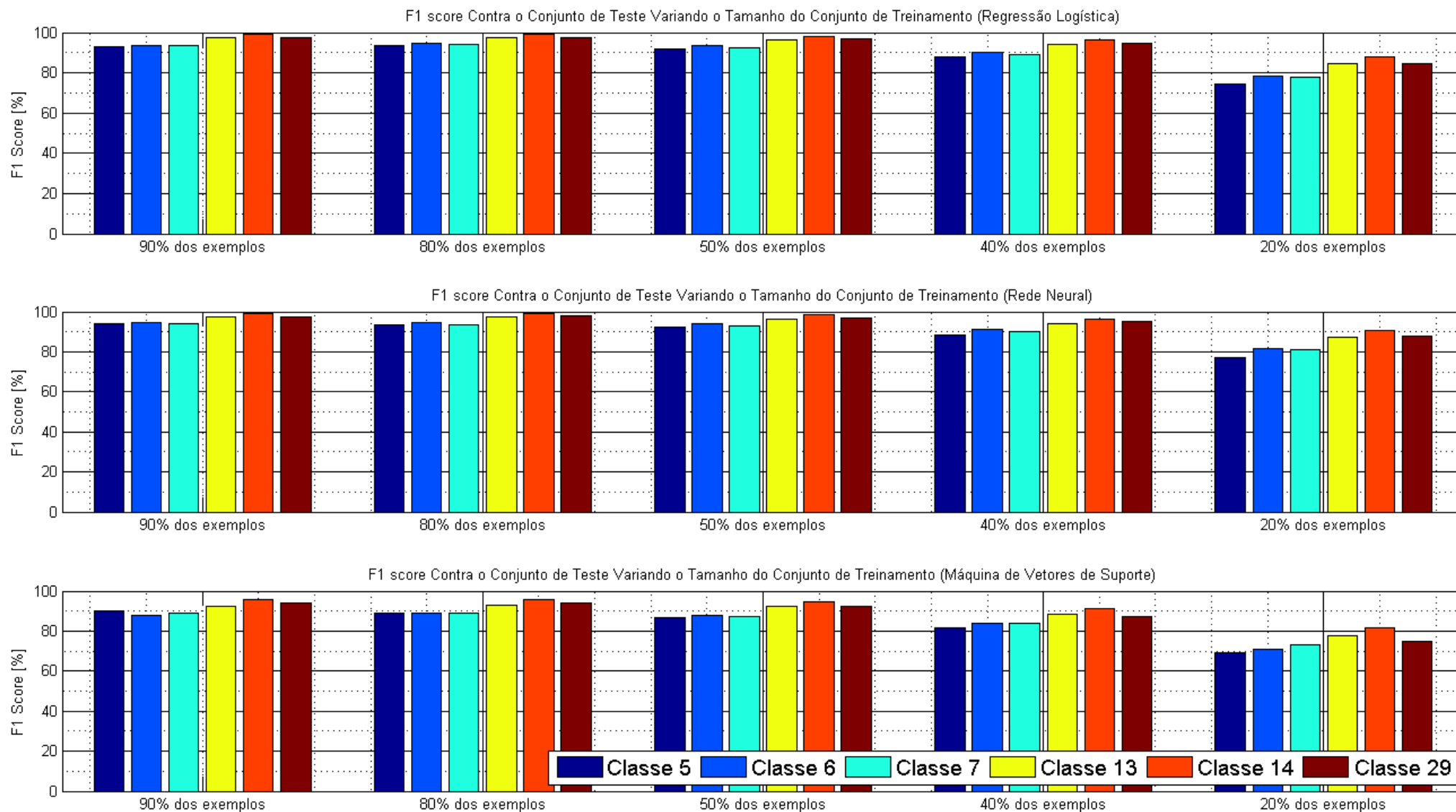
Nota-se, observando os *F1 scores* de cada classe para cada caso de teste, pela Figura 8.7, que todos os três modelos de aprendizado considerados apresentam maiores dificuldades em classificar as classes 5, 6 e 7, que correspondem, respectivamente, às poses de abdução dos dedos, dedos flexionados em punho e extração do dedo indicador (veja

Figura 8.6: Taxa de acerto variando o tamanho do conjunto de treinamento.



a Figura 3.3). Como se deveria esperar, essas dificuldades aumentam com a utilização de menores conjuntos de treinamento.

Figura 8.7: F1 score variando o tamanho do conjunto de treinamento.



9 CONCLUSÕES

O objetivo deste trabalho, como mencionado anteriormente, é a avaliação do desempenho de três algoritmos de classificação (regressão logística, redes neurais e máquina de vetores de suporte) para conjuntos variáveis de eletrodos, para conjuntos variáveis de características e para tamanhos de conjunto de treinamento distintos.

Mostrou-se, na Seção 8.1, que a quantidade de eletrodos utilizados interfere, diretamente, na taxa de acerto média de cada um dos modelos de aprendizado considerados. Quanto menor o número de eletrodos considerados, menor é a taxa de acerto média do sistema contra o conjunto de teste. Além disso, encontrou-se evidência de que o posicionamento de eletrodos em músculos selecionados de acordo com os movimentos a serem identificados pode melhorar a taxa de acerto dos algoritmos para esses movimentos.

A Seção 8.2 deixa claro que nenhuma das características descritas na Seção 4.5 é insubstituível na presença das outras. Destaca-se, também, que sistemas que consideram apenas características pertencentes ao domínio da frequência apresentam maiores taxas de acerto que sistemas que consideram somente características pertencentes ao domínio do tempo. Além disso, esta seção mostra que uma das características consideradas - a média aritmética do sinal no domínio do tempo - é, sozinha, incapaz de classificar os movimentos da base de dados considerada (quando comparando as métricas de qualidade de classificação deste caso aos outros semelhantes) e que algumas características - como a obliquidade (no domínio da frequência) e, principalmente, o centroide espectral - obtêm, sozinhas, taxas de acerto médias próximas à taxa de acerto do sistema completo, considerando todas as características descritas na Seção 4.5.

Na Seção 8.3, ficou claro que a taxa de acerto média dos modelos de aprendizado diminui com a redução do tamanho do conjunto de treinamento e que, entre os movimentos considerados, aqueles que apresentam as menores taxas de acerto médias são os representados pelas classes 5, 6 e 7 (ou, respectivamente, abdução dos dedos, dedos flexionados em punho e extração do dedo indicador).

Finalmente, nas Seções 8.1, 8.2 e 8.3, mostrou-se que os modelos de regressão logística e de rede neural apresentam taxas de acerto médias muito próximas, enquanto que o modelo de máquina de vetores de suporte apresenta taxas de acerto inferiores às taxas dos outros modelos.

Tendo em vista os resultados apresentados no Capítulo 8, sugere-se, como configuração básica para a classificação das poses da base de dados NinaPro consideradas neste trabalho, o uso de, no mínimo, 8 eletrodos (que podem ser posicionados de forma igualmente espaçada em volta do antebraço, independentemente de músculos individuais). Afim de atingir as melhores taxas de acerto, os sinais vindos destes eletrodos podem ser classificados por um modelo de regressão logística ou de rede neural, utilizando, no mínimo, como características de entrada para a classificação, o centroide espectral dos sinais.

REFERÊNCIAS

ATKINS, D. J.; HEARD, D. C.; DONOVAN, W. H. Epidemiologic Overview of Individuals with Upper-Limb Loss and their Reported Research Priorities. **JPO: Journal of Prosthetics and Orthotics**, [S.l.], v.8, n.1, p.2–11, 1996.

ATZORI, M.; GIJSBERTS, A.; HEYNEN, S.; HAGER, A.-G. M.; DERIAZ, O.; VAN DER SMAGT, P.; CASTELLINI, C.; CAPUTO, B.; MULLER, H. Building the NINA-PRO database: a resource for the biorobotics community. In: BIOMEDICAL ROBOTICS AND BIOMECHATRONICS (BIOROB), 2012 4TH IEEE RAS & EMBS INTERNATIONAL CONFERENCE ON. **Proceedings...** [S.l.: s.n.], 2012. p.1258–1265.

BASMAJIAN, J. V.; DE LUCA, C. Muscles alive. **Baltimore/USA**, [S.l.], 1985.

BELTRAMINI, L. **Elementos de Histologia e Anatomo-fisiologia Humana**. [S.l.]: IFSC, 1999.

BISHOP, C. M. **Pattern Recognition and Machine Learning**. [S.l.]: Springer, 2006.

BOSCHMANN, A.; KAUFMANN, P.; PLATZNER, M.; WINKLER, M. Towards multi-movement hand prostheses: combining adaptive classification with high precision sockets. In: EUROPEAN CONFERENCE TECHNICALLY ASSISTED REHABILITATION, 2. **Proceedings** [S.l.: s.n.], 2009.

BRESADOLA, M. Carlo Matteucci and the legacy of Luigi Galvani. **Archives Italiennes de Biologie**, [S.l.], v.149, n.Supplement, p.3–9, 2011.

FAVIEIRO, G. W. **Controle de uma Prótese Experimental do Segmento Mão-Braço por Sinais Mioelétricos e Redes Neurais Artificiais**. 2009. — Universidade Federal do Rio Grande do Sul.

FAVIEIRO, G. W. **Desenvolvimento de um Sistema Neuro-Fuzzy para Análise de Sinais Mioelétricos do Segmento Mão-Braço**. 2012. — Universidade Federal do Rio Grande do Sul.

FAVIEIRO, G. W.; BALBINOT, A. Adaptive neuro-fuzzy logic analysis based on myoelectric signals for multifunction prosthesis control. In: ENGINEERING IN MEDICINE AND BIOLOGY SOCIETY, EMBC, 2011 ANNUAL INTERNATIONAL CONFERENCE OF THE IEEE. **Proceedings...** [S.l.: s.n.], 2011. p.7888–7891.

FAVIEIRO, G. W.; BALBINOT, A.; BARRETO, M. M. Decoding arm movements by myoelectric signals and artificial neural networks. In: BIOSIGNALS AND BIOROBOTICS CONFERENCE (BRC), 2011 ISSNIP. **Proceedings...** [S.l.: s.n.], 2011. p.1–6.

FAWCETT, T. An introduction to ROC analysis. **Pattern recognition letters**, [S.l.], v.27, n.8, p.861–874, 2006.

GIJSBERTS, A.; ATZORI, M.; CASTELLINI, C.; MULLER, H.; CAPUTO, B. The Movement Error Rate for Evaluation of Machine Learning Methods for sEMG-based Hand Movement Classification. **IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society**, [S.l.], 2014.

GLETTE, K.; GRUBER, T.; KAUFMANN, P.; TORRESEN, J.; SICK, B.; PLATZNER, M. Comparing evolvable hardware to conventional classifiers for electromyographic prosthetic hand control. In: ADAPTIVE HARDWARE AND SYSTEMS, 2008. AHS'08. NASA/ESA CONFERENCE ON. **Proceedings...** [S.l.: s.n.], 2008. p.32–39.

GUYTON, A. C.; HALL, J. E. **Textbook of Medical Physiology**. 12th.ed. [S.l.]: Elsevier Health Sciences, 2010.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The Elements of Statistical Learning**. 2nd.ed. [S.l.]: Springer, 2009.

LIN, C. J.; HSU, C.; CHANG, C. A practical guide to support vector classification. **National Taiwan U., www.csie.ntu.edu.tw/cjlin/papers/guide/guide.pdf**, [S.l.], 2003.

MAZURKIEWICZ, P. Automatic Segmentation of EMG Signals Based on Wavelet Representation. In: **Computer Recognition Systems 2**. [S.l.]: Springer, 2007. p.589–595.

NILSON, C. d. P. **Aquisição, Processamento de Sinais Mioelétricos e Máquina de Vetores de Suporte para Caracterização de Movimentos do Segmento Mão-Braço**. 2014. — Universidade Federal do Rio Grande do Sul.

ORTOLAN, R. L. **Estudo e Avaliação de Técnicas de Processamento do Sinal Mioelétrico para o Controle de Sistemas de Reabilitação**. 2002. — Universidade de São Paulo.

PEARCE, J. Emil Heinrich Du Bois-Reymond (1818–96). **Journal of Neurology, Neurosurgery & Psychiatry**, [S.l.], v.71, n.5, p.620–620, 2001.

PEETERS, G. A Large set of audio features for sound description (similarity and classification) in the CUIDADO project. , [S.l.], 2004.

POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. , [S.l.], 2011.

RICCIOTTI, A. C. D. **Utilização de Wavelets no Processamento de Sinais EMG**. 2006. — Universidade Federal de Uberlândia.

SCHRÖDER, G.; THIELE, M.; LEHNER, W. Setting Goals and Choosing Metrics for Recommender System Evaluations. In: SECOND WORKSHOP ON USER-CENTRIC EVALUATION OF RECOMMENDER SYSTEMS AND THEIR INTERFACES (UCERSTI 2). **Proceedings** [S.l.: s.n.], 2011.

TANG, X.; LIU, Y.; LV, C.; SUN, D. Hand motion classification using a multi-channel surface electromyography sensor. **Sensors**, [S.l.], v.12, n.2, p.1130–1147, 2012.

TORTORA, G. J.; DERRICKSON, B. **Principles of Anatomy & Physiology**. 14th.ed. [S.l.]: Wiley, 2013.

WHITTAKER, E. T. History of the Theories of Ether and Electricity from the Age of Descartes to the Close of the Nineteenth Century. **London and Dublin**, [S.l.], 1910.

ANEXO A FUNÇÃO DE MINIMIZAÇÃO POR CARL EDWARD RASMUSSEN

```

function [X, fX, i] = fmincg(f, X, options, P1, P2, P3, P4, P5)
% Minimize a continuous differentiable multivariate function. Starting point
% is given by "X" (D by 1), and the function named in the string "f", must
% return a function value and a vector of partial derivatives. The Polack-
% Ribiere flavour of conjugate gradients is used to compute search directions,
% and a line search using quadratic and cubic polynomial approximations and the
% Wolfe-Powell stopping criteria is used together with the slope ratio method
% for guessing initial step sizes. Additionally a bunch of checks are made to
% make sure that exploration is taking place and that extrapolation will not
% be unboundedly large. The "length" gives the length of the run: if it is
% positive, it gives the maximum number of line searches, if negative its
% absolute gives the maximum allowed number of function evaluations. You can
% (optionally) give "length" a second component, which will indicate the
% reduction in function value to be expected in the first line-search (defaults
% to 1.0). The function returns when either its length is up, or if no further
% progress can be made (ie, we are at a minimum, or so close that due to
% numerical problems, we cannot get any closer). If the function terminates
% within a few iterations, it could be an indication that the function value
% and derivatives are not consistent (ie, there may be a bug in the
% implementation of your "f" function). The function returns the found
% solution "X", a vector of function values "fX" indicating the progress made
% and "i" the number of iterations (line searches or function evaluations,
% depending on the sign of "length") used.
%
% Usage: [X, fX, i] = fmincg(f, X, options, P1, P2, P3, P4, P5)
%
% See also: checkgrad
%
% Copyright (C) 2001 and 2002 by Carl Edward Rasmussen. Date 2002-02-13
%
%
% (C) Copyright 1999, 2000 & 2001, Carl Edward Rasmussen
%
% Permission is granted for anyone to copy, use, or modify these
% programs and accompanying documents for purposes of research or
% education, provided this copyright notice is retained, and note is
% made of any changes that have been made.
%
% These programs and documents are distributed without any warranty,
% express or implied. As the programs were written for research
% purposes only, they have not been tested to the degree that would be
% advisable in any important application. All use of these programs is

```

```

% entirely at the user's own risk.
%
% [ml-class] Changes Made:
% 1) Function name and argument specifications
% 2) Output display
%
% Read options
if exist('options', 'var') && ~isempty(options) && isfield(options, 'MaxIter')
    length = options.MaxIter;
else
    length = 100;
end

RHO = 0.01; % a bunch of constants for line searches
SIG = 0.5; % RHO and SIG are the constants in the Wolfe-Powell conditions
INT = 0.1; % don't reevaluate within 0.1 of the limit of the current bracket
EXT = 3.0; % extrapolate maximum 3 times the current bracket
MAX = 20; % max 20 function evaluations per line search
RATIO = 100; % maximum allowed slope ratio

argstr = ['feval(f, X']; % compose string used to call function
for i = 1:(nargin - 3)
    argstr = [argstr, ',P', int2str(i)];
end
argstr = [argstr, ')'];

if max(size(length)) == 2, red=length(2); length=length(1); else red=1; end
S=['Iteration_'];

i = 0; % zero the run length counter
ls_failed = 0; % no previous line search has failed
fX = [];
[f1 df1] = eval(argstr); % get function value and gradient
i = i + (length < 0); % count epochs?!
s = -df1; % search direction is steepest
d1 = -s*s; % this is the slope
z1 = red/(1-d1); % initial step is red/(|s|+1)

while i < abs(length) % while not finished
    i = i + (length > 0); % count iterations?!

    X0 = X; f0 = f1; df0 = df1; % make a copy of current values
    X = X + z1*s; % begin line search
    [f2 df2] = eval(argstr);
    i = i + (length < 0); % count epochs?!
    d2 = df2*s;
    f3 = f1; d3 = d1; z3 = -z1; % initialize point 3 equal to point 1
    if length > 0, M = MAX; else M = min(MAX, -length-i); end
    success = 0; limit = -1; % initialize quantities
    while 1
        while ((f2 > f1+z1*RHO*d1) | (d2 > -SIG*d1)) & (M > 0)
            limit = z1; % tighten the bracket
            if f2 > f1
                z2 = z3 - (0.5*d3*z3*z3)/(d3*z3+f2-f3); % quadratic fit
            else
                A = 6*(f2-f3)/z3+3*(d2+d3); % cubic fit
            end
        end
    end

```

```

    B = 3*(f3-f2)-z3*(d3+2*d2);
    z2 = (sqrt(B*B-A*d2*z3*z3)-B)/A;           % numerical error possible - ok!
end
if isnan(z2) | isinf(z2)
    z2 = z3/2;                               % if we had a numerical problem then bisect
end
z2 = max(min(z2, INT*z3),(1-INT)*z3); % don't accept too close to limits
z1 = z1 + z2;                               % update the step
X = X + z2*s;
[f2 df2] = eval(argstr);
M = M - 1; i = i + (length < 0);           % count epochs?!
d2 = df2'*s;
z3 = z3-z2;                               % z3 is now relative to the location of z2
end
if f2 > f1+z1*RHO*d1 | d2 > -SIG*d1
    break;                                   % this is a failure
elseif d2 > SIG*d1
    success = 1; break;                     % success
elseif M == 0
    break;                                   % failure
end
A = 6*(f2-f3)/z3+3*(d2+d3);                % make cubic extrapolation
B = 3*(f3-f2)-z3*(d3+2*d2);
z2 = -d2*z3*z3/(B+sqrt(B*B-A*d2*z3*z3));   % num. error possible - ok!
if ~isreal(z2) | isnan(z2) | isinf(z2) | z2 < 0 % num prob or wrong sign?
    if limit < -0.5                          % if we have no upper limit
        z2 = z1 * (EXT-1);                   % the extrapolate the maximum amount
    else
        z2 = (limit-z1)/2;                   % otherwise bisect
    end
elseif (limit > -0.5) & (z2+z1 > limit)      % extrapolation beyond max?
        z2 = (limit-z1)/2;                   % bisect
elseif (limit < -0.5) & (z2+z1 > z1*EXT)    % extrapolation beyond limit
        z2 = z1*(EXT-1.0);                  % set to extrapolation limit
elseif z2 < -z3*INT
        z2 = -z3*INT;
elseif (limit > -0.5) & (z2 < (limit-z1)*(1.0-INT)) % too close to limit?
        z2 = (limit-z1)*(1.0-INT);
end
f3 = f2; d3 = d2; z3 = -z2;                % set point 3 equal to point 2
z1 = z1 + z2; X = X + z2*s;                 % update current estimates
[f2 df2] = eval(argstr);
M = M - 1; i = i + (length < 0);           % count epochs?!
d2 = df2'*s;
end
% end of line search

if success                                % if line search succeeded
    f1 = f2; fX = [fX' f1]';
    %fprintf('%s %4i | Cost: %4.6e\r', S, i, f1);
    s = (df2'*df2-df1'*df2)/(df1'*df1)*s - df2; % Polack-Ribiere direction
    tmp = df1; df1 = df2; df2 = tmp;        % swap derivatives
    d2 = df1'*s;
    if d2 > 0                               % new slope must be negative
        s = -df1;                            % otherwise use steepest direction
        d2 = -s'*s;
    end
    z1 = z1 * min(RATIO, d1/(d2-realmin));   % slope ratio but max RATIO
    d1 = d2;

```

```

    ls_failed = 0;                                     % this line search did not fail
else
    X = X0; f1 = f0; df1 = df0; % restore point from before failed line search
    if ls_failed | i > abs(length) % line search failed twice in a row
        break; % or we ran out of time, so we give up
    end
    tmp = df1; df1 = df2; df2 = tmp; % swap derivatives
    s = -df1; % try steepest
    d1 = -s'*s;
    z1 = 1/(1-d1);
    ls_failed = 1; % this line search failed
end
if exist('OCTAVE_VERSION')
    fflush(stdout);
end
end
%fprintf('\n');

```