

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

FABIANO ROMERO DE SOUZA BADARACO

**CLASSIFICAÇÃO BINÁRIA DE MOVIMENTOS
IMAGINÁRIOS DA FLEXÃO DA MÃO CAPTURADOS POR
ELETROENCEFALOGRAFIA ATRAVÉS DE UM SISTEMA
NEURO-*FUZZY***

Porto Alegre

2019

FABIANO ROMERO DE SOUZA BADARACO

**CLASSIFICAÇÃO BINÁRIA DE MOVIMENTOS
IMAGINÁRIOS DA FLEXÃO DA MÃO CAPTURADOS POR
ELETROENCEFALOGRAFIA ATRAVÉS DE UM SISTEMA
NEURO-*FUZZY***

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Engenharia Elétrica, da Universidade Federal do Rio Grande do Sul, como parte dos requisitos para a obtenção do título de Mestre em Engenharia Elétrica.

Área de concentração: Controle e Automação
– Instrumentação.

ORIENTADOR: Professor Dr. Alexandre Balbinot

Porto Alegre

2019

FABIANO ROMERO DE SOUZA BADARACO

**CLASSIFICAÇÃO BINÁRIA DE MOVIMENTOS
IMAGINÁRIOS DA FLEXÃO DA MÃO CAPTURADOS POR
ELETROENCEFALOGRAFIA ATRAVÉS DE UM SISTEMA
NEURO-FUZZY**

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia Elétrica e aprovada em sua forma final pelo Orientador e pela Banca Examinadora.

Orientador: _____

Prof. Dr. Alexandre Balbinot, UFRGS - PPGEE

Doutor pela UFRGS, Porto Alegre

Banca Examinadora:

Prof. Dr. Henrique Takachi Moriya, PPGEE – USP

Doutor em Engenharia Elétrica pela Universidade de São Paulo, São Paulo, Brasil.

Prof. Dr. Raphael Martins Brum, DELET – UFRGS

Doutor em Microeletrônica pelo Univesrité Montpellier 2 – Science et Technique, França.

Prof. Dr. Valner João Brusamarello, DELET – UFRGS

Doutor em Ciência e Engenharia de Materiais pela Universidade Federal de Santa Catarina, Brasil.

Coordenador do PPGEE: _____

Prof. Dr. João Manoel Gomes da Silva Junior

Porto Alegre, setembro 2019.

DEDICATÓRIA

“Sozinho vamos mais rápido, juntos vamos mais longe.”

“Ubuntu – sou porque somos”.

Autores desconhecidos, provérbios africanos.

Dedico esse trabalho aos meus pais e por sua exigência para comigo, pois devido a isso, me fizeram ser uma pessoa exigente, mola propulsora que me permitiu ir em busca dos meus sonhos.

A minha mãe, agradeço por ter me ensinado a importância de estudar, devido à falta que o conhecimento lhe fazia.

Ao meu pai, agradeço o ensinamento de ter paciência e de entender o tempo que é necessário para tudo acontecer. Infelizmente lamento que ele não possa estar presente fisicamente para presenciar a conclusão deste trabalho, mas acredito que estará presente espiritualmente para tal.

AGRADECIMENTOS

Ao Programa de Pós-Graduação em Engenharia Elétrica, PPGEE, pela oportunidade de desenvolver este trabalho na área de *BCI*, área de pesquisa a qual me dediquei.

Aos colegas do PPGEE pelo seu auxílio nas tarefas desenvolvidas durante o curso e apoio no aprendizado mútuo através da troca de conhecimento ao longo das disciplinas e do desenvolvimento do trabalho.

Deixo aqui um agradecimento ao professor Alexandre Balbinot, muito obrigado pelo apoio e paciência no desenvolvimento deste trabalho.

A CAPES, pela provisão da bolsa de mestrado, pois sem ela não seria possível a dedicação que dispus a este trabalho.

RESUMO

Este trabalho consiste na avaliação do método *Neuro-Fuzzy* associado ao *Subtractive Clustering* como classificador de sinais de eletroencefalografia para sistemas *Brain Computer Interface* (BCI), utilizando a combinação de três diferentes características extraídas deste sinal, durante a imaginação do movimento de flexão da mão direita ou esquerda. Para isto é avaliado a ocorrência dos fenômenos ERD e ERS, com o objetivo de determinar em qual segmentação no tempo devem ser extraídas as características do mesmo. Os sinais de Eletroencefalografia avaliados são provenientes do banco de dados do *BCI Competition* e de um ensaio adquirido no Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS. O método proposto é testado e comparado com outro método de aprendizagem de máquina denominado *Multi Layer Percetron* (MLP) treinado com o algoritmo Levenberg-Marquardt (LM), com o objetivo de conhecer qual o método e a combinação de características apresenta a maior taxa de acerto. Em sequência, estas taxas de acerto são comparadas com as taxas de acerto de outros trabalhos da área. Os métodos propostos neste trabalho alcançaram uma taxa de acerto máxima de $92,6 \pm 1,1\%$ para a ANFIS SC e $87,9 \pm 1,5\%$ para a MLP LM. Por fim, é apresentada uma análise estatística dos resultados alcançados.

Palavras-chave: Processamento de Sinais de EEG, Movimento Imaginário, ERD/ERS, ANFIS, MLP.

ABSTRACT

This work evaluates the Neuro-Fuzzy method associated with Subtractive Clustering as an electroencephalographic signal classifier for Brain Computer Interface (BCI) systems, using the combination of three different features extracted from this signal, during the imagination of the flexion movement of the right or left hand. For this, the occurrence of the ERD and ERS phenomena is evaluated, in order to determine in which segmentation in time the characteristics of the same should be extracted. The electroencephalography signals evaluated come from the BCI Competition database and from a trial acquired at the UFRGS Electroelectronic & Biosignal Instrumentation Laboratory. The chosen method is tested and compared with another machine learning method called Multi Layer Percetron (MLP) trained with the Levenberg-Marquardt (LM) algorithm, in order to know which method and combination of features has the highest rate. As a result, these hit rates are compared to hit rates from other jobs in the area. The methods proposed in this study reached a maximum accuracy of $92.6 \pm 1.1\%$ for ANFIS SC and $87.9 \pm 1.5\%$ for MLP LM. Finally, a statistical analysis of the results achieved is presented.

Keywords: EEG Signal Processing, Imaginary Movement, ERD/ERS, ANFIS MLP.

LISTA DE FIGURAS

Figura 1 -	Áreas funcionais do Telencéfalo.....	18
Figura 2 -	Mapa Funcional das áreas sensorial e motora.....	19
Figura 3 -	Distribuição dos eletrodos no escalpo conforme sistema 10-20 (para 21 eletrodos) e suas respectivas nomenclaturas.....	21
Figura 4 -	Passo a passo para a detecção de ERD e ERS.....	22
Figura 5 -	Diagrama de Blocos de um típico Sistema de Inferência <i>Fuzzy</i>	24
Figura 6 -	Sistema ANFIS com duas entradas.....	25
Figura 7 -	Diagrama de Blocos do sistema proposto.....	34
Figura 8 -	Temporização das trilhas.....	35
Figura 9 -	Posição dos eletrodos.....	36
Figura 10 -	Temporização das trilhas do sujeito O3 do <i>BCI Competition IIIb</i> (2005).....	36
Figura 11 -	<i>Basket Paradigm</i> , utilizado no experimento realizado nos sujeitos S4 e X11.....	37
Figura 12 -	Posições dos eletrodos no <i>Emotiv Epoc System®</i>	37
Figura 13 -	Aparato utilizado.....	38
Figura 14 -	Foto de um dos ensaios.....	39
Figura 15 -	Fluxograma do procedimento de geração do estímulo.....	40
Figura 16 -	Índice de Lateralização calculado a partir do sinal do sujeito S1 do <i>BCI Competition II</i>	41
Figura 17 -	Ocorrência dos fenômenos ERD e ERS a partir do sinal do sujeito S1 do <i>BCI Competition II</i>	42
Figura 18 -	Diagrama de blocos de um exemplo de entrada com a combinação de duas características (Energia e RMS).....	43
Figura 19 -	A) Agrupamento de dados de entrada; B) Centros de <i>cluster</i> identificados pelo algoritmo <i>Subtractive Clustering</i>	44
Figura 20 -	Funções de Pertinência geradas pelo algoritmo <i>Subtractive Clustering</i>	45
Figura 21 -	Exemplo de topologia ANFIS gerada pela técnica <i>Subtractive Clustering</i>	45
Figura 22 -	Diagrama de blocos da rede neural proposta.....	46
Figura 23 -	Ocorrência do fenômeno ERD e ERS nos sujeitos S1 e O3.....	48
Figura 24 -	Ocorrência do fenômeno ERD e ERS nos sujeitos S4 e X11.....	49
Figura 25 -	Ocorrência dos fenômenos ERD e ERS no sujeito V1 do Laboratório do IEE & Biosinais da UFRGS.....	49
Figura 26 -	Performance da RNA MLP LM para os sujeitos do <i>BCI Competition</i> ..	50
Figura 27 -	Performance do sujeito V1 da base de dados do IEE & Biosinais da (UFRGS).....	51
Figura 28 -	Performance da ANFIS SC para os sujeitos S1 e O3.....	52
Figura 29 -	Performance da ANFIS SC para os sujeitos S4 e X11.....	52
Figura 30 -	Performance da ANFIS SC para os sujeitos V1.....	53
Figura 31 -	Análise dos resíduos do Projeto de Experimentos utilizando os dados do <i>BCI Competition</i>	58
Figura 32 -	Análise das Taxas de Acerto das Combinações de Características.....	60
Figura 33 -	Análise das Taxas de Acerto do Fator Controlável Classe.....	61
Figura 34 -	Análise das Taxas de Acerto do Fator Controlável Sujeito.....	62

Figura 35 -	Análise das Taxas de Acerto do Fator Controlável Método.....	63
Figura 36 -	Interação entre os Fatores Controláveis Características e Método.....	64
Figura 37 -	Interação entre os Fatores Controláveis Sujeito e Característica.....	65
Figura 38 -	Interação entre os Fatores Controláveis Método e Classe.....	66
Figura 39 -	Interação entre os Fatores Controláveis Sujeito e Classe.....	67
Figura 40 -	Interação entre os Fatores Controláveis Características e Classe.....	68
Figura 41 -	Interação entre os Fatores Controláveis Método, Sujeito e Classe.....	69
Figura 42	Neurônio artificial proposto por McCulloch e Pitts.....	79
Figura 43	Gráfico da Interação Método x Sujeito x Classe.....	83
Figura 44	Gráfico da Interação Sujeito x Característica x Classe.....	83
Figura 45	Gráfico da Interação Método x Sujeito x Característica x Classe.....	84

LISTA DE TABELAS

Tabela 1 -	Publicações selecionadas para a classificação do EEG.....	32
Tabela 2 -	Informações dos bancos de dados utilizados.....	39
Tabela 3 -	Projeto de Experimento realizado, baseado nos resultados de classificação.....	47
Tabela 4 -	Segmentação onde ocorrem os fenômenos ERD e ERS.....	50
Tabela 5 -	Taxas de acerto para o sujeito S1 utilizando o algoritmo MLP LM.....	54
Tabela 6 -	Taxas de acerto para o sujeito O3 utilizando o algoritmo MLP LM.....	54
Tabela 7 -	Taxas de acerto para o sujeito S4 utilizando o algoritmo MLP LM.....	54
Tabela 8 -	Taxas de acerto para o sujeito X11 utilizando o algoritmo MLP LM.....	54
Tabela 9 -	Taxas de acerto para o sujeito S1 utilizando o algoritmo ANFIS SC.....	55
Tabela 10 -	Taxas de acerto para o sujeito O3 utilizando o algoritmo ANFIS SC.....	55
Tabela 11 -	Taxas de acerto para o sujeito S4 utilizando o algoritmo ANFIS SC.....	55
Tabela 12 -	Taxas de acerto para o sujeito X11 utilizando o algoritmo ANFIS SC.....	55
Tabela 13 -	<i>Ranking</i> dos trabalhos pesquisados.....	56
Tabela 14 -	Taxas de acerto máximas para os ensaios do IEE UFRGS.....	57

LISTA DE ABREVIATURAS

AFNN	<i>Adaptative Fuzzy Neural Network</i>
ANFIS	<i>Adaptative Neuro Fuzzy Inference System</i>
BCI	<i>Brain Computer Interface</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CSP	<i>Common Spatial Pattern</i>
DOE	<i>Design of Experiments</i>
DSLQVQ	<i>Districtive Sensitive Learning Vector Quantization</i>
ECoG	Electrocorticograma
EEG	Eletroencefalograma
EMD	<i>Empirical Mode Decomposition</i>
ERD	<i>Event Related Desynchronization</i>
ERS	<i>Event Related Synchronization</i>
FCM	<i>Fuzzy C-Means</i>
FIS	Sistema de Inferência <i>Fuzzy</i>
HMI	Interface Homem Máquina
ICA	<i>Independent Component Analysis</i>
IEE	Instrumentação Eletroeletrônica
IIR	<i>Infinite Impulse Response</i>
IL	Índice de Lateralização
IT2FLF-ANFIS	<i>Interval Type-2 Fuzzy Logic ANFIS fusion</i>
KNN	<i>K-Nearest Neighbor</i>
LDA	<i>Linear Discriminant Analysis</i>
LM	Levenberg-Marquardt

LVQ	<i>Learning Vector Quantization</i>
MAV	<i>Mean Absolute Value</i>
MLP	Perceptron Multicamada
NB	<i>Nayve Bayes</i>
OVA	<i>One versus All</i>
OVO	<i>One versus One</i>
PCA	<i>Principal Component Analysis</i>
PSD	Densidade Espectral de Potência
PW	Periodograma de <i>Welch</i>
RBF	<i>Radial Basis Functions</i>
RMS	<i>Root Mean Square</i>
RNA	Redes Neurais Artificiais
RNN	Rede Neural Recorrente
SAFIS	<i>Sequential Adaptative Fuzzy Inference System</i>
SC	<i>Subtractive Clustering</i>
STDF	<i>Spatio-Temporal Discrepancy Feature</i>
SVM	<i>Support Vector Machine</i>

SUMÁRIO

1.	INTRODUÇÃO.....	15
2.	REVISÃO DA LITERATURA.....	17
2.1	PRINCÍPIOS SOBRE ELETROENCEFALOGRAFIA (EEG) E POTENCIAIS RELACIONADOS A EVENTOS.....	20
2.2	SISTEMA DE INFERÊNCIA FUZZY E NEURO-FUZZY.....	23
2.3	SUBSTRUCTIVE CLUSTERING.....	27
2.4	TRABALHOS RELACIONADOS À CLASSIFICAÇÃO DE SINAIS EEG/ERD-ERS.....	28
2.5	ABORDAGEM SIMPLIFICADA SOBRE PRÉ-PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS.....	32
3.0	METODOLOGIA.....	34
3.1	BANCO DE DADOS BCI <i>COMPETITION</i>.....	34
3.2	BASE DE DADOS GERADA.....	37
3.3	PRÉ-PROCESSAMENTO.....	41
3.4	EXTRAÇÃO DE CARACTERÍSTICAS.....	42
3.5	PROCESSAMENTO E CLASSIFICAÇÃO.....	43
3.6	MÉTODO DE COMPARAÇÃO E ANÁLISE ESTATÍSTICA.....	45
3.6.1	Análise Estatística: DOE.....	46
4.	RESULTADOS E DISCUSSÕES.....	48
4.1	PROJETO DE EXPERIMENTOS.....	57
4.2	VALIDAÇÃO DO PROJETO DE EXPERIMENTOS.....	58
4.2.1	Resultados do Projeto Experimentos.....	59
4.2.2	Análise do Fator Controlável Combinação de Características.....	60

4.2.3	Análise do Fator Controlável Classe.....	61
4.2.4	Análise do Fator Controlável Sujeito.....	62
4.2.5	Análise do Fator Controlável Método.....	63
4.2.6	Análise da Interação entre os Fatores Controláveis Método e Características.....	64
4.2.7	Análise da Interação entre os Fatores Controláveis Sujeito e Características.....	65
4.2.8	Análise da Interação entre os Fatores Controláveis Método e Classe.....	66
4.2.9	Análise da Interação entre os Fatores Controláveis Sujeito e Classe.....	67
4.2.10	Análise da Interação entre os Fatores Controláveis Característica e Classe...	68
4.2.11	Análise da Interação entre os Fatores Controláveis Método, Sujeito e Característica.....	69
5.	CONCLUSÕES.....	71
6.	PROPOSTA DE TRABALHOS FUTUROS.....	73
	REFERÊNCIAS BIBLIOGRÁFICAS.....	74
	APÊNDICE A.....	78
	ANEXO A.....	83

1. INTRODUÇÃO

O registro do primeiro estudo sobre Eletroencefalografia (EEG) está datado em meados de 1929, na Universidade de Jena pelo psiquiatra austríaco Hans Berger, utilizando um galvanômetro Siemens de bobina dupla, com um eletrodo bipolar e um canal com derivações fronto-occipitais, identificando o ritmo α (alfa) como principal componente dos sinais de Eletroencefalografia (SANEI; CHAMBERS, 2007).

Interfaces cérebro-computador (*Brain Control Interface, BCI*) surgiram como uma nova alternativa para proporcionar às pessoas que sofrem deficiências motoras parciais ou completas, com um canal de comunicação não muscular para transmitir mensagens ou comandos para o mundo externo, seja ele para controlar dispositivos como um mouse, uma cadeira de rodas ou até mesmo um exoesqueleto. Consequentemente, esses sistemas podem ajudar a melhorar a qualidade de vida do usuário, dando mais independência e autonomia, enquanto constitui uma nova ferramenta de pesquisa para a compreensão do Cérebro (ANTELIS *et al.*, 2017).

Em geral as interfaces cérebro computador têm por objetivo auxiliar três principais grupos de pessoas: o primeiro inclui pacientes que perderam todo o controle do movimento por encontrarem-se na fase final da esclerose lateral amiotrófica ou sofrem com distúrbios cerebrais graves. O segundo grupo é composto por pessoas que encontram-se completamente paralisadas, mas ainda conseguem controlar alguns movimentos como o dos olhos, por exemplo. E, o terceiro, composto por pessoas capazes de mover e controlar o nervo muscular essencial ou a fala (MOHAMMADPOUR *et al.* 2016).

Desenvolver uma inteligência artificial que possa auxiliar pessoas com limitações das mais diversas formas foi o que gerou o interesse por essa área e motivou o desenvolvimento do estudo aqui apresentado.

Dentre as diversas técnicas disponíveis para a detecção de atividade cerebral, tais como: Magnetoencefalografia, Ressonância Magnética Funcional, Potencial Evocado, entre outras, esta pesquisa fará uso do Eletroencefalograma (EEG) na detecção de padrões de potenciais elétricos cerebrais relacionados ao movimento das mãos (MIRVAZIRI; MOBARAKEH, 2017).

Neste contexto, a presente pesquisa tem como objetivo principal analisar uma rede Neuro-Fuzzy como classificador de sinais de Eletroencefalografia, visando reconhecer dados destes sinais relacionados à execução de movimentos imaginários da flexão de mãos. Para isso, o Periodograma de Welch, a Energia e o *Root Mean Square* (RMS) de um segmento do

sinal foram utilizados como característica para alimentar o classificador. Como objetivo específico visa-se comparar as taxas de acerto da *Neuro-Fuzzy* com uma rede MLP (método clássico) e, por fim, definir qual combinação das características extraídas do sinal proporciona a melhor taxa de acerto.

2. REVISÃO DA LITERATURA

Este capítulo apresenta uma breve revisão da literatura, iniciando-se com a descrição do Cérebro Humano, sua divisão e composição, com foco principal no Córtex Motor. Depois, aborda princípios sobre Eletroencefalografia (EEG) e Potenciais Relacionados a Eventos, e descreve o sistema de inferência Neuro-*Fuzzy*. Também são apresentados trabalhos relacionados à classificação de EEG, selecionados em bases de dados para o período de cinco anos. Por fim, apresenta uma introdução aos procedimentos de pré-processamento e de extração de características.

O Cérebro Humano é dividido em Encéfalo e Cerebelo, sendo o Encéfalo subdividido em Tronco Encefálico, Diencefalo e Telencefalo. O Tronco Encefálico é contínuo à Medula Espinhal e consiste em Bulbo, Ponte e Mesencefalo. Acima do Tronco Encefálico encontra-se o Diencefalo, constituído por Tálamo, Hipotálamo e Glândula Pineal. Apoiado sobre o Diencefalo e o Tronco Encefálico está o Telencefalo, formando a maior massa do Encéfalo (TORTORA, 2000; SOBOTTA, 2000).

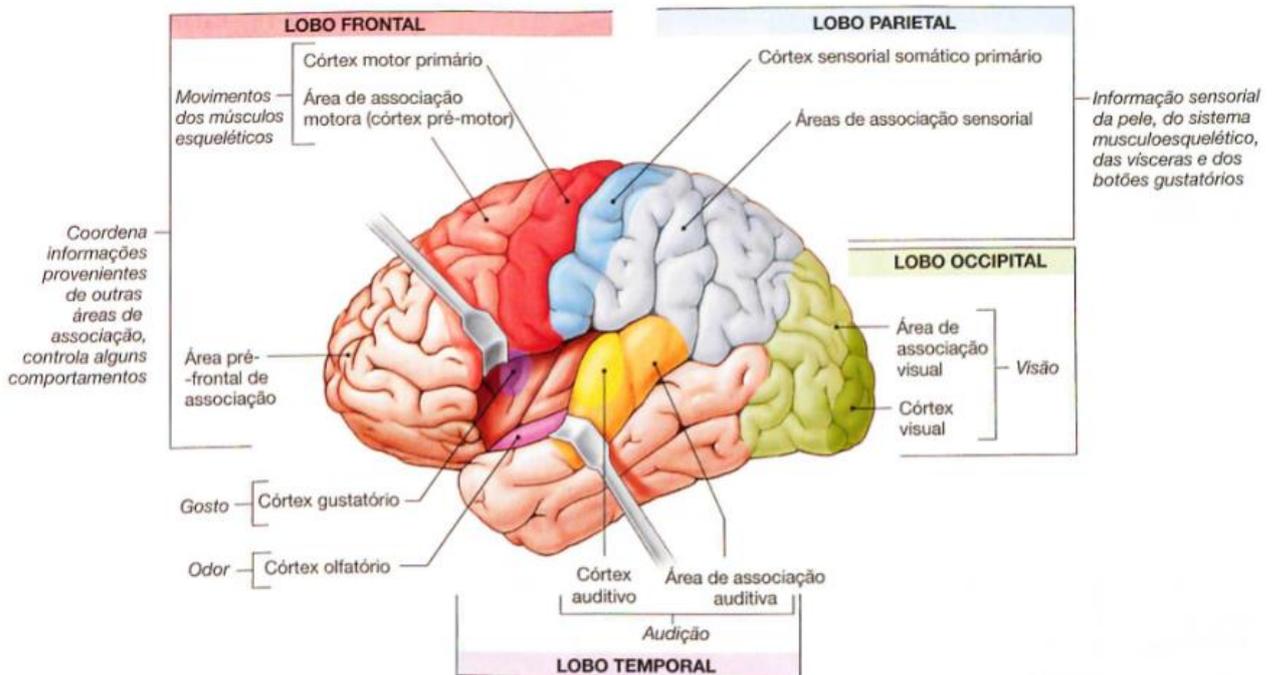
Portanto, no que se relaciona a este trabalho, o interesse é destinado à compreensão do funcionamento do Córtex Cerebral, em especial às áreas relacionadas aos movimentos. Sendo assim, segundo Sobotta (2000), o Córtex Cerebral pode ser dividido nas seguintes áreas:

- áreas sensitivas: a informação sensitiva do Córtex Cerebral flui principalmente da metade posterior dos hemisférios cerebrais, para regiões atrás dos sulcos centrais. A área somatossensitiva primária é posterior ao sulco central de cada hemisfério cerebral. Esta área é responsável pela percepção de dor, tato, temperatura e cócegas;
- áreas motoras: a resposta motora do Córtex Cerebral flui principalmente da parte anterior de cada hemisfério. Entre as áreas motoras mais importantes, estão a área motora primária e a área da fala de Broca. Cada região na área motora primária controla as contrações voluntárias de músculos específicos, no lado oposto do corpo. A área da fala de Broca está localizada no Lobo Frontal, próximo ao sulco cerebral lateral. A fala e a compreensão da linguagem são atividades complexas que envolvem várias áreas sensitivas, de associação e motoras do Córtex;

- áreas de associação: as áreas de associação do Telencéfalo consistem em grandes áreas dos Lobos Occipital, Parietal e Temporal, e do Lobo Frontal anterior às áreas motoras. Tratos conectam as áreas de associação umas às outras. A área de associação somatossensitiva, posterior à área somatossensitiva primária, integra e interpreta sensações somáticas, como forma exata e textura de um objeto, por exemplo.

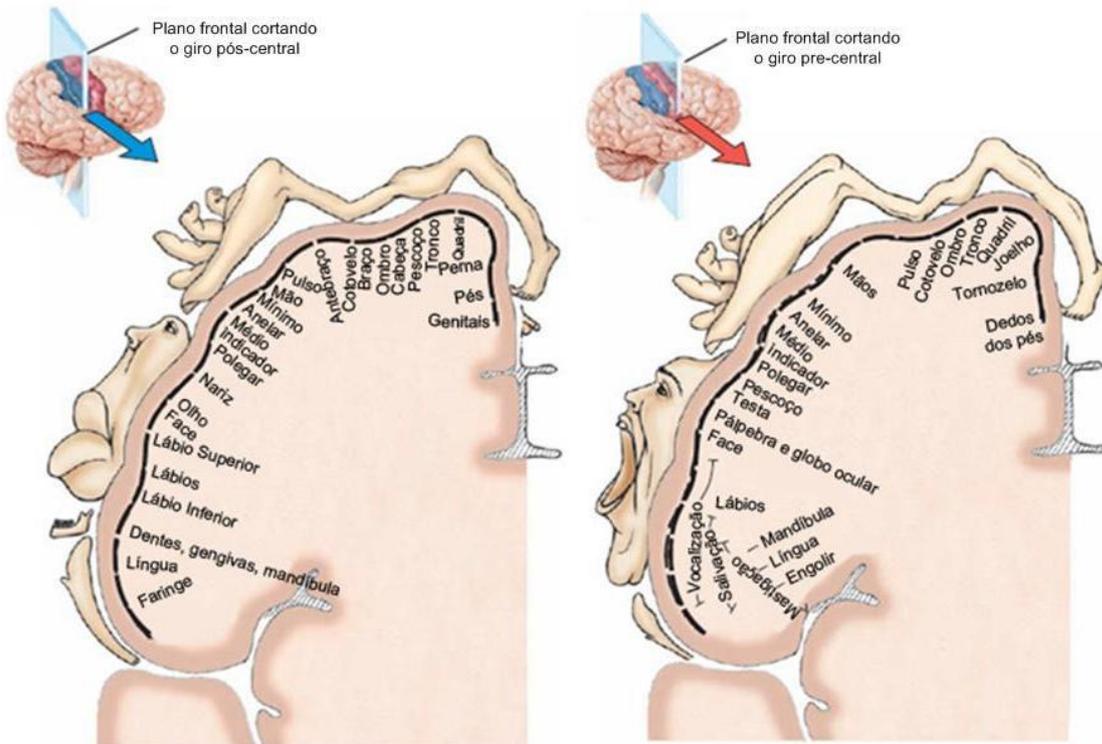
Estas áreas de associação estão apresentadas na Figura 1 e na Figura 2; nesta última, está representada graficamente a distribuição destas áreas.

Figura 1 - Áreas funcionais do Telencéfalo.



Fonte: Silverthorn (2010, p.292).

Figura 2 - Mapa Funcional das áreas sensorial e motora.



Fonte: Sobotta (2000).

De acordo com Guyton e Hall (2003) e Sobotta (2000), muitas atividades do sistema nervoso se iniciam pelas experiências sensoriais que excitam os receptores sensoriais, sejam eles receptores visuais, receptores auditivos, receptores táteis na superfície do corpo ou qualquer outro tipo de receptor. Essa experiência pode provocar uma reação imediata, ou mesmo ser memorizada para determinar futuras reações no organismo. Por fim, o objetivo mais importante do sistema nervoso é controlar as diversas atividades do corpo, tais como a contração dos músculos esqueléticos apropriados, contração da musculatura lisa dos órgãos internos e da secreção de substâncias químicas pelas glândulas exócrinas e endócrinas que agem em diversas partes do corpo. Tais atividades denominam-se funções motoras do sistema nervoso, os músculos e glândulas são denominados efetores, pois são as estruturas anatômicas que verdadeiramente executam as funções comandadas pelos sinais nervosos.

Segundo Guyton e Hall (2003) impulsos nervosos, ou potenciais de ação é como denomina-se a forma como o sistema nervoso central transmite a informação propagada através da sucessão de neurônios, um após o outro. Tais potenciais de ação são ocasionados devido à troca de íons através da membrana de um neurônio; sua transmissão ocorre através

de seu axônio, esta membrana despolariza-se e, após, volta ao normal. Este evento tem a duração entre 5 a 10 milissegundos (SANEI; CHAMBERS, 2007; KROPOTOV, 2009).

2.1 PRINCÍPIOS SOBRE ELETROENCEFALOGRAFIA (EEG) E POTENCIAIS RELACIONADOS A EVENTOS

No Século XIX, o psiquiatra Hans Berger descreveu qualitativamente a maneira como funciona o Cérebro em animais sem anestesia; com esta primeira análise iniciaram-se os estudos sobre a Eletroencefalografia. As alterações químicas neuronais geram campos elétricos no Córtex Cerebral, correntes elétricas fluem no Córtex Cerebral devido a esta atividade dos conjuntos neuronais e esses estímulos podem ser monitorados, por exemplo, no couro cabeludo com a utilização de eletrodos não invasivos. Quando este sinal é consequência de um estímulo externo, seja ele visual ou auditivo, as correntes no Córtex geram um campo eletromagnético que pode ser medido através de um equipamento denominado de EEG. Estas correntes observadas no escalpo durante diferentes atividades classificam-se em seis tipos de ritmos, denominados de ondas α , μ , θ , β , δ , e γ (SILVA, 2014). Na detecção de movimento imaginário da região motora os principais sinais encontram-se nas frequências α e μ (8 a 13Hz) e β (14 a 26Hz), com amplitudes inferiores a $50\mu\text{V}$ e $30\mu\text{V}$. Estas frequências também são chamadas de ritmos sensório-motores (BATTACHARYYA *et al.*, 2015).

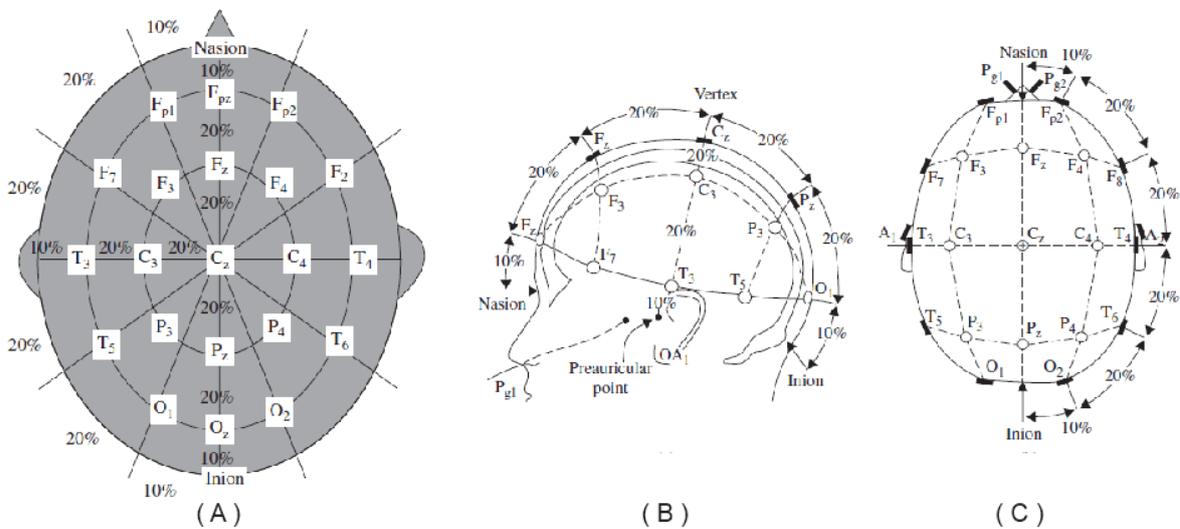
As ondas α e μ são encontradas na mesma banda de frequência e com amplitudes semelhantes, porém com origens diferentes: enquanto o ritmo α se origina na região Occipital, o ritmo μ se origina no Córtex Motor (SANEI; CHAMBERS, 2007; KANDEL *et al.*, 2000).

As ondas β estão relacionadas ao ritmo habitual de vigília do Cérebro, associado ao pensamento ativo, atenção ativa, foco no mundo externo ou solução de problemas concretos. Estas atividades rítmicas podem ser encontradas principalmente nas regiões frontais e central do Córtex Cerebral (SANEI; CHAMBERS, 2007)

Os sinais de Eletroencefalografia podem ser capturados através de eletrodos invasivos, implantados cirurgicamente no Córtex Cerebral, o Electrocorticograma (ECoG), ou através de eletrodos não invasivos fixados sobre o couro cabeludo (escalpo), que podem ser fixados livremente no escalpo, ou mesmo fixos a uma touca. O sistema de aquisição não invasivo possui uma relação sinal-ruído pior que o sistema invasivo, no entanto, é largamente utilizado em estudos com humanos devido aos riscos envolvidos na intervenção cirúrgica no Córtex Cerebral (MOBARAKEH *et al.*, 2017).

Segundo Jasper (1958), tipicamente, a distribuição dos eletrodos no escalpo respeita a convenção internacional denominada Sistema 10-20. Este sistema distribui os eletrodos entre pontos de interesse de medição, utilizando 10 ou 20% desta distância conforme apresentado na Figura 3 (JASPER, 1958).

Figura 3 - Distribuição dos eletrodos no escalpo conforme sistema 10-20 (para 21 eletrodos) e suas respectivas nomenclaturas.



Fonte: Adaptado de Jasper (1958).

Ao planejar um movimento, imaginar, ou mesmo executar tal movimento, o Córtex Cerebral tem a capacidade de prever este acontecimento. Em torno de 1 a 1,5 segundos antes da execução do movimento, há uma alteração nos ritmos sensório-motores, em especial nas ondas μ , mas também pode ser observado nas ondas β . Tal fenômeno é denominado de *Event Related Desynchronization* (ERD), e ocorre de forma mais evidente no lado oposto ao que o indivíduo tem a intenção de movimentar (no chamado lado contralateral). No entanto, no mesmo lado em que o indivíduo tem a intenção de executar o movimento (chamado lado ipsilateral) é possível perceber um fenômeno inverso, ou seja, ocorre um aumento na energia e este fenômeno é denominado de *Event Related Synchronization* (ERS) (NAM *et al.*, 2011).

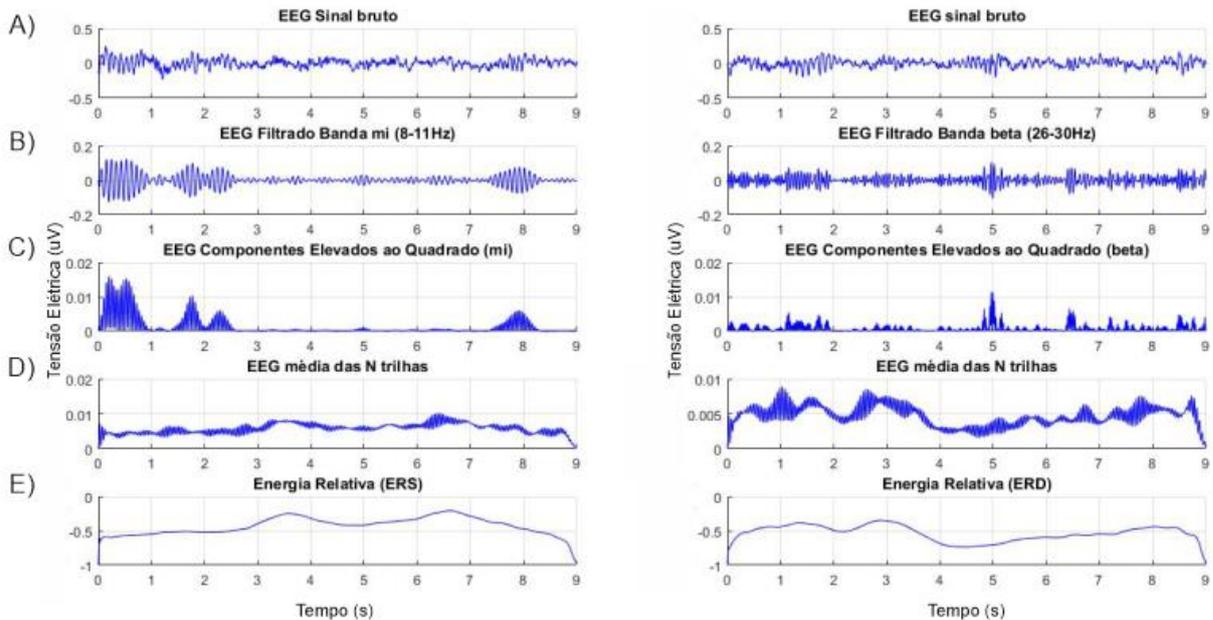
Com o objetivo de detectar ambos os fenômenos, pode-se utilizar o método proposto por Pfurtscheller e Silva (1999), no qual avalia-se a energia relativa entre o período de estímulo e o período de referência com base nos seguintes passos:

1. filtrar os sinais separando os ritmos μ e β ;
2. elevar as componentes ao quadrado obtendo a potência do sinal;
3. calcular a média entre as trilhas que possuem o mesmo estímulo;
4. calcular a potência relativa conforme apresentado na Equação (1).

$$ERD\% = \frac{A - R}{R} * 100 \quad (1)$$

onde A representa a potência do sinal durante o estímulo e R é a potência do sinal durante o período de referência. Apenas para ilustrar o método exposto, a Figura 4 apresenta um exemplo deste método aplicado a um sinal de EEG bruto (Figura 4A). Este mesmo sinal é filtrado nas bandas de frequência μ e β conforme apresentado na Figura 4B. Na Figura 4C estão apresentados os componentes do EEG elevados ao quadrado. Na Figura 4D é apresentado o sinal médio da energia ao longo do tempo. E, por fim, na Figura 4E, a energia relativa calculada a partir da Equação (1).

Figura 4 - Passo a passo para a detecção de ERD e ERS utilizando os sinais do sujeito S1 do BCI Competition II.



Fonte: Autor.

É importante lembrar que a modulação destes ritmos pode variar para cada indivíduo, podendo impactar na detecção de ERD e ERS. Além disso, a repetição das tarefas pelo

indivíduo pode aumentar a taxa de acerto na identificação destes fenômenos (BAMDADIAN *et al.*, 2014; NAKAYASHIKI *et al.*, 2014).

Como os fenômenos (ERD e ERS) apresentam características de lateralização, isto possibilita a definição de um parâmetro relacionado à diferença de energia entre os dois hemisférios Cerebrais, denominado de Índice de Lateralização (IL). Este parâmetro é calculado como a média da diferença de potência entre dois eletrodos posicionados um de cada lado do Córtex para cada classe de movimento, conforme indicado pela Equação (2).

$$IL = \frac{(P_{Ch1} - P_{Ch2})\text{Movimento Esquerdo} + (P_{Ch2} - P_{Ch1})\text{Movimento Direito}}{2} \quad (2)$$

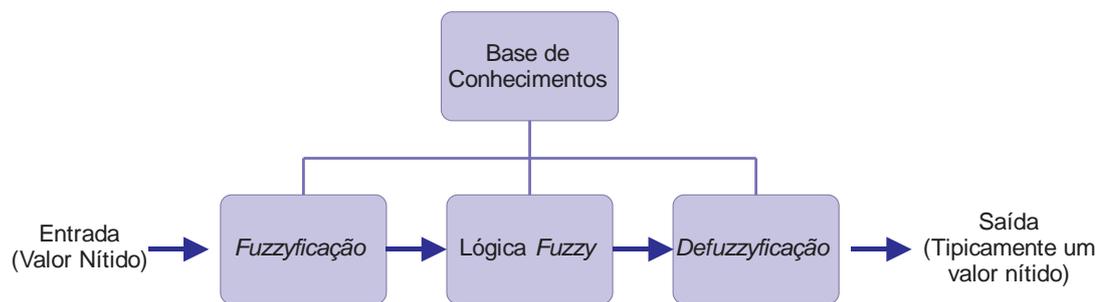
onde IL é o índice de lateralização, $Ch1$ representa o eletrodo posicionado no córtex motor esquerdo, $Ch2$ o eletrodo posicionado no córtex motor direito, P_{Ch1} e P_{Ch2} são as respectivas potências dos sinais determinados dos sinais captados pelos eletrodos posicionados no lado esquerdo e no lado direito.

2.2 SISTEMAS DE INFERÊNCIA FUZZY E NEURO-FUZZY

De acordo com Goble (2001) e Willi *et al.* (2015) o raciocínio matemático sobre conjuntos possibilita um raciocínio lógico para a maioria dos problemas do cotidiano dos indivíduos. Por exemplo, considere um conjunto de dados denominado de X e um determinado Y que pode ou não pertencer ao conjunto X . Este correspondente valor Y poderia representar um horário do dia ou da noite com a utilização do atributo bivalente. Porém, como definir se o horário pertence ao entardecer com a correspondente lógica bivalente?

Zadeh, em 1965, propôs um sistema inteligente capaz de combinar técnicas e metodologias de diferentes fontes e de solucionar problemas que necessitam de uma resposta em uma região de transição (como o exemplo anterior: o entardecer, entre o dia e a noite) entre o verdadeiro e o falso, este sistema foi denominado como Sistema *Fuzzy*. O procedimento de raciocínio *Fuzzy*, ou mais comumente chamado de Sistema de Inferência *Fuzzy* (FIS), de forma genérica, é composto pelas três etapas principais representadas no diagrama de blocos da Figura 5 (ZADEH, 1965).

Figura 5 – Diagrama de Blocos de um típico Sistema de Inferência *Fuzzy*.



Fonte: Adaptado livremente de Jang (1993).

A primeira etapa do FIS é denominada *Fuzzyficação*. Nesta etapa, as variáveis de entrada (normalmente com valores nítidos, ou seja, não *Fuzzy*) são convertidas em uma escala de valores conforme as funções de pertinência definidas em um determinado universo de discurso.

A segunda etapa do FIS é chamada de *Lógica* ou *Regras Fuzzy* e possui regras normalmente definidas por um conjunto de pessoas (denominadas de especialistas) com base no conjunto de conhecimentos das mesmas. Essas regras são semelhantes às decisões tomadas pelos indivíduos no cotidiano e são definidas com variáveis linguísticas, onde cada regra aplicada (antecedente), tem por consequência uma saída (consequente), portanto, cada regra agregada define a saída com a utilização dos operadores *Fuzzy*. Existem diversos sistemas de inferência como, por exemplo: o *DE-Induced Fuzzy-Neural* (SAHA *et al.*, 2015), o SAFIS (*Sequential Adaptive Fuzzy Inference System*) (RONG *et al.*, 2006), entre outros, mas os mais utilizados ainda são o Sistema Mamdani (MAMDANI; ASSILAN, 1975) e o Sistema de Inferência Sugeno^{1*} (TAKAGI; SUGENO, 1985).

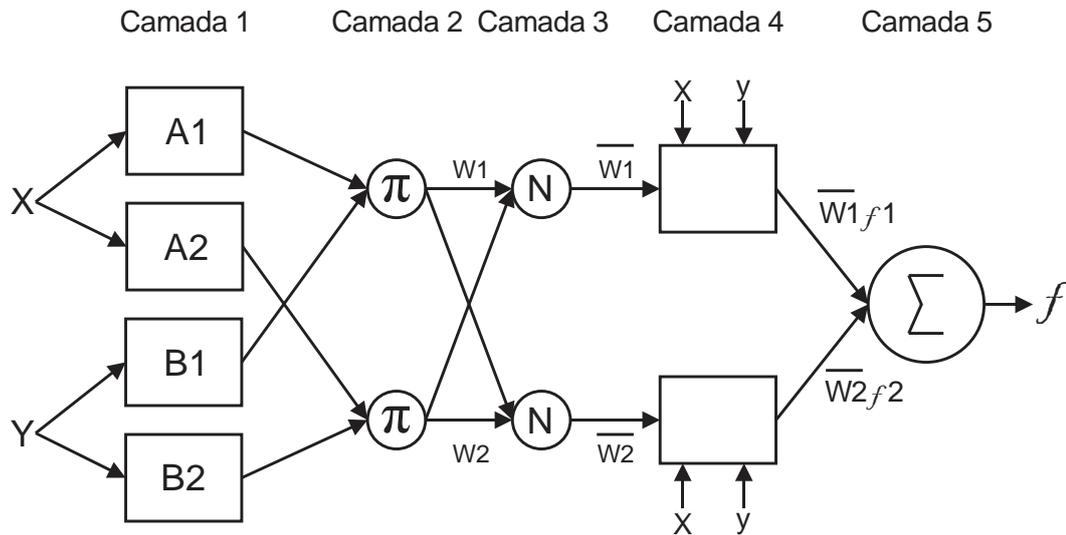
A terceira etapa do FIS é chamada de *Defuzzyficação* e é responsável por transformar os resultados das funções de pertinência e regras do FIS em valores nítidos na saída.

Como o Sistema Neuro-*Fuzzy* empregado neste trabalho é o ANFIS (JANG, 1993) que só utiliza o Sistema do Tipo Sugeno, não serão apresentados os demais tipos de FIS existentes. O ANFIS (*Adaptive Neuro Fuzzy Inference System*) é uma classe de redes adaptativas que tem uma funcionalidade similar a um Sistema de Inferência *Fuzzy* e é capaz

¹ O controlador proposto por Takagi e Sugeno (1985) consiste em uma simplificação do controlador desenvolvido por Mandani, onde o consequente de cada regra é definido como uma função das variáveis

de gerar uma base de regras *Fuzzy* e funções de pertinência automaticamente (JANG, 1993). Como exemplo a Figura 6 apresenta uma típica topologia ANFIS com duas entradas.

Figura 6 – Sistema ANFIS com duas entradas.



Fonte: Adaptado livremente de Jang (1993).

Com base na Figura 6 e no trabalho de Jang (1997), pode-se descrever as camadas da rede ANFIS da seguinte forma:

- camada 1: é composta por nós adaptativos, onde a saída de cada nó é o valor de pertinência dado pela função de pertinência de entrada. Esta camada é responsável pelo ajuste dos parâmetros das funções de pertinência e a função deste nó está representada na Equação (3):

$$\begin{aligned} O_{1,i} &= \mu_{A_i}(x), & \text{para } i = 1,2 & \text{ ou} \\ O_{1,i} &= \mu_{B_{i-2}}(y), & \text{para } i = 3,4 \end{aligned} \quad (3)$$

onde x ou y é a entrada do nó i e A_i ou B_{i-2} é o rótulo linguístico associado ao nó, de modo que $O_{1,i}$ é o nível de pertinência para um conjunto *Fuzzy* A ($A = A_1, A_2, B_1$ e B_2) especificando o grau de cada entrada (x ou y), que satisfaz o quantificador A ;

- camada 2: é composta por nós não adaptativos, onde cada nó é responsável pela multiplicação dos sinais de entrada e entrega destes sinais multiplicados ao

próximo nó, representando “a força de cada regra”. Cada nó desta camada é um nó fixo rotulado de π , cuja saída é o produto de todos os sinais de entrada, representado pela Equação (4):

$$O_{2,i} = \bar{w}_l = \mu_{A_i}(x)\mu_{B_i}(x), \quad i = 1, 2 \quad (4)$$

Nesta camada 2, cada nó de saída caracteriza a força de disparo de uma regra;

- camada 3: cujos nós adaptativos determinam o peso de cada regra e que podem ser representados pela Equação (5):

$$O_{3,i} = \bar{w}_l = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad (5)$$

Este nó é chamado de forças de disparo normalizado;

- camada 4: é composta por nós adaptativos que aplicam os pesos aos parâmetros de entrada para determinar o consequente de aplicação de cada regra e são representados pela Equação (6):

$$O_{4,i} = \bar{w}_l f_i = \bar{w}_l (p_i x + q_i y + r_i), \quad (6)$$

onde \bar{w} é a força de disparo normalizada da camada 3 e $(p_i, q_i$ e $r_i)$ são o conjunto de parâmetros desse nó. Estes parâmetros são chamados de parâmetros das consequências;

- camada 5: é composta por nós fixos que determinam o valor de saída a partir da soma de todos os sinais de entrada definido pela Equação (7):

$$Saída\ final = O_{5,1} = \sum_i \bar{w}_l f_i = \frac{\sum_i w_i f_i}{\sum_i w_i}, \quad (7)$$

Estas camadas possibilitam a construção de uma rede adaptativa com funcionalidade equivalente a um modelo de Inferência *Fuzzy* Sugeno.

Com base na arquitetura da ANFIS apresentada na Figura 6, é observado que enquanto os valores dos parâmetros das premissas estão fixos, a saída final pode ser expressa como na Equação (8):

$$f = \begin{cases} \frac{w_1}{w_1+w_2} f_1 + \frac{w_2}{w_1+w_2} f_2 \\ \overline{w}_l(p_1x + q_1y + r_1) + \overline{w}_l(p_2x + q_2y + r_2) \\ (\overline{w}_l x)p_1 + (\overline{w}_l y)q_1 + (\overline{w}_l)r_1 + (\overline{w}_l x)p_2 + (\overline{w}_l y)q_2 + (\overline{w}_l)r_2 \end{cases} \quad (8)$$

A Equação (8) é linear nos parâmetros de consequência p_1, q_1, r_1, p_2, q_2 e r_2 . A partir dessa observação pode-se afirmar que:

- S é igual ao conjunto total dos parâmetros;
- S1 é igual ao conjunto dos parâmetros das premissas (não linear);
- S2 é igual ao conjunto dos parâmetros das consequências (linear).

Portanto, o algoritmo híbrido aqui descrito e proposto por Jang (1993), pode ser aplicado diretamente. De forma geral, no passo para frente (*feedforward*) do algoritmo híbrido, as saídas dos nós vão para frente em direção a camada 4 e os parâmetros das consequências são identificados pelo método dos mínimos quadrados. No passo para trás (*backpropagation*) o erro é propagado para trás e os parâmetros das premissas são atualizados pelo método do gradiente descendente (JANG, 1997).

Para caracterização e organização de dados, algoritmos de agrupamento de dados (*Clustering*, que trata-se de um método de aprendizado não supervisionado) são extensivamente utilizados, mas não são interessantes para a construção de modelos. O agrupamento de dados separa dados em diferentes grupos conforme uma métrica de similaridade. Esta técnica pode ser utilizada para o modelamento inicial de uma rede *Fuzzy* como apresentado, por exemplo, em Hsu (2014). A técnica de agrupamento de dados considera que entradas similares em um sistema alvo devem ser modeladas para saídas similares e que tais pares de entrada-saída, são agrupados em *clusters* do conjunto de dados de treinamento.

2.3 SUBSTRUCTIVE CLUSTERING

Chiu, em 1994, propôs a técnica denominada *subtractive clustering*, que considera os pontos dos dados candidatos para centros dos *clusters*, dessa forma o processamento é proporcional ao número de pontos de dados e independente da dimensão do problema em consideração.

Considere, por exemplo, um grupo com n pontos de dados (x_1, \dots, x_n) em um espaço de dimensão M , assumindo que os dados foram normalizados em um hipercubo e cada ponto

de dado é candidato para o centro do *cluster*, a medida de densidade em cada ponto x_i pode ser definida pela Equação (9):

$$D_i = \sum_{j=1}^n \exp\left(\frac{\|x_i - x_j\|^2}{(r_a/2)^2}\right) \quad (9)$$

onde r_a é uma constante positiva. Portanto, um ponto terá uma alta densidade se houverem muitos pontos vizinhos, logo o raio r_a definirá a vizinhança e os pontos fora desse raio de vizinhança, não contribuirão, ou contribuirão muito pouco para a medida de densidade.

Agora que a medida de densidade foi calculada para todos os pontos, o ponto que apresentar a maior densidade é selecionado para ser o centro do primeiro *cluster*. Considere x_{c1} o ponto selecionado e D_{c1} sua medida de densidade, a densidade medida para cada ponto x_i pode ser determinada pela Equação (10):

$$D_i = D_i - D_{c1} \exp\left(\frac{\|x_i - x_{c1}\|^2}{(r_b/2)^2}\right) \quad (10)$$

em que r_b é uma constante positiva. Logo, os pontos mais próximos do primeiro centro de *cluster* x_{c1} , apresentarão suas medidas de densidade reduzidas, evitando que esses pontos sejam escolhidos como o próximo centro. Conforme sugerido por Chiu (1994), em geral r_b deve ser $1,5 r_a$.

Por fim, depois que cada ponto teve sua densidade determinada, o próximo centro x_{c2} é definido e todas as medidas de densidade são determinadas novamente e assim sucessivamente até que um número satisfatório de *clusters* seja definido. Ao aplicar essa técnica em um conjunto de dados de entrada-saída cada centro de *cluster* representará um agrupamento do sistema a ser modelado. Este modelo será utilizado como centro das premissas das regras *Fuzzy* em um modelo Sugeno de ordem zero.

2.4 TRABALHOS RELACIONADOS À CLASSIFICAÇÃO DE SINAIS DE EEG/ERD-ERS

A engenharia biomédica vem estudando sistemas *BCI* há muitos anos e recentemente tem utilizado variadas técnicas de inteligência artificial para classificar os sinais de eletroencefalografia com as mais diferentes formas de extrações de características e também diferentes objetivos de controle como, por exemplo, o controle de um braço robótico (BHATTACHARYYA *et al.*, 2015) ou o controle de uma cadeira de rodas (CARRA, 2012).

O estudo apresentado por Luo *et al.* (2019) mostra uma técnica de extração de características denominada *Spatio-Temporal Discrepancy Feature* (STDF) e utilizando o classificador *Support Vector Machines* (SVM) associado a *Time Series Voting*, para identificar movimento imaginário da mão direita e esquerda a partir dos fenômenos ERD e ERS. Estas técnicas foram avaliadas no *BCI Competition IV* e alcançaram sua melhor taxa de acerto com 85,0%.

O trabalho proposto por Mohammadpour *et al.* (2016) compara dois classificadores de base, *Decision Stump* (DS) e *K-Nearest Neighbors* (KNN), associados aos métodos: *Bagging* – que trata-se de um agrupamento que utiliza o método de amostragem de *bootstrap* para criar conjuntos a partir de dados de treinamento, ajustando cada classificador em um subconjunto aleatório do conjunto de treinamento; *Boosting* – que é um método como o *Bagging*, dependente da amostragem do conjunto de dados de treinamento. O objetivo dessa abordagem é produzir uma coleção de classificadores para aumentar a precisão da classificação. Ao contrário do *Bagging*, a amostragem nesse método é adaptativa e é afetada pelo desempenho do classificador treinado anteriormente; *AdaBoost* – que trata-se de um método metaheurístico no aprendizado de máquina. Seu objetivo é melhorar a precisão e exatidão do algoritmo de aprendizado combinando classificadores fracos. E são extraídas quatro características do sinal: *Autoregressive Parameter* (AR), *Discrete Wavelet Transform* (DWT), *Common Spatial Patterns* (CSP) e *Power Spectrum Density* (PSD). As características são apresentadas aos classificadores individualmente e o classificador que obteve a melhor taxa de acerto, foi o KNN associado ao método *Adaboost*, com a característica CSP, 73,4%.

O estudo realizado por Bhattacharyya *et al.* (2015) apresenta a interface de controle com o Cérebro (BCI) como uma Interface Homem Máquina (HMI), para controlar um braço robótico denominado *Jaco Robot Arm* utilizando a *Emotiv Epoc System* de forma *on line*, aplicando o algoritmo *Fuzzy Type-2²* aplicado a sinais de Eletroencefalografia. Com esse classificador foram avaliadas duas técnicas de validação: *One Versus All* (OVA) com a abordagem de construir N diferentes classes, onde cada classificador corresponde a cada classe individual e *One Versus One* (OVO) que utiliza um classificador binário para separação de dados. A avaliação da taxa de acerto foi determinada com um índice de confiança de 95,0%. A performance da combinação das técnicas utilizadas OVA-IT2FLF-ANFIS e OVO-IT2FLF-ANFIS foi comparada com suas concorrentes: *Linear Discriminant*

² Conjuntos e sistemas *Fuzzy Type 2* generalizam conjuntos e sistemas *Fuzzy* padrão *Type 1*, para que maiores incertezas possam ser tratadas.

Analysis (LDA), *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM) e *Naive Bayes* (NB) através do Teste de Friedman e obtiveram uma taxa de acerto maior que suas concorrentes, 88,9% e 90,9%, respectivamente. Os resultados *on line* e *off line* demonstraram que a técnica IT2FLF-ANFIS direcionada a detecção multi-classe é eficiente.

O estudo apresentado por Saha *et al.* (2015) descreve uma técnica de classificação utilizando *Neuro-Fuzzy*, associando Rede Neural Recorrente (RNN) e ANFIS. A rede recorrente é utilizada para padronizar a classificação, e a definição da rede começa utilizando a função de Lyapunov para verificar sua dinâmica e estabilidade. Então, uma função de Lyapunov com múltiplos mínimos bem definidos, é escolhida para o armazenamento de padrões e a dinâmica é derivada pela acumulação das condições para a estabilidade de Lyapunov. O pré-processamento utiliza técnicas de filtros *Infinite Impulse Response* (IIR) conhecidas para separar três bandas, sendo que os filtros utilizados foram *Butterworth*, *Chebyshev* e Elíptico, e as bandas utilizadas foram θ (3-7Hz), α (7-13Hz) e β (13-30Hz). As técnicas utilizadas neste trabalho conseguiram uma taxa de acerto de 95,5%.

O estudo desenvolvido por Hsu (2014) avaliou técnicas de classificação de EEG utilizando o banco de dados do *BCI Competition* e na etapa de pré-processamento utilizou a Transformada Wavelet Contínua (*Continuos Wavelet Transform*) associada a outras técnicas de filtro. As técnicas avaliadas foram: *Adaptative Fuzzy Neural Network* (AFNN), *Support Vector Machine* (SVM) e *Linear Discriminant Analysis* (LDA) e as taxas de acerto foram: 79,3% (LDA), 81,6% (MLP), 82,3% (AFNN).

O trabalho apresentado por Wei *et al.* (2014) utilizou *Fuzzy C-means* (FCM) como classificador de sinais de eletroencefalografia. Alguns algoritmos são utilizados para classificar sinais de eletroencefalografia (EEG) como, por exemplo: SVM, LDA e MLP, mas todos estes necessitam que ocorra um número considerável de treinamentos antecipadamente, o que não é necessário com a técnica apresentada neste trabalho. Para a extração de características foi utilizado *Empirical Mode Decomposicion* (EMD), para decompor os sinais de EEG em diferentes componentes de frequência e extrair o valor absoluto médio (MAV). O banco de dados utilizado para este trabalho foi o *BCI Competition II* (2003). A taxa de acerto da técnica alvo deste trabalho (EMD MAV FCM) alcançou 78,0%, abaixo dos 80,0% da EMD MAV SVM utilizada para comparação, mas acima das outras duas técnicas avaliadas, a MAV FCM (61,0%) e a MAV SVM (70,0%).

Machado e Balbinot (2014) apresentaram um estudo utilizando o classificador *Naive Bayes* (NB), utilizando como pré-processamento um estimador espectral denominado Periodograma de *Welch* (PW), associado ao *Common Spatial Pattern* (CSP). Estas

características foram apresentadas aos classificadores *Naive Bayes* (NB) e *Linear Discriminant Analysis* (LDA), observando-se a ocorrência do fenômeno ERD e ERS e em qual janela de tempo obteve-se a maior taxa de acerto do classificador, alcançando uma taxa de acerto de 70,0%.

Lin *et al.* (2013) utilizaram a técnica *Fuzzy C-Means* para identificar os fenômenos ERD e ERS em sinais de Eletroencefalografia (EEG), utilizando o banco de dados *BCI Competition II*. As bandas de interesse do movimento imaginário foram filtradas através de filtros Butterworth e então apresentadas ao classificador, denominado *Fuzzy C-means*. Este classificador apresentou a maior taxa de assertividade durante o instante de tempo de 3,5~7 segundos, (82,8%).

Utilizando o banco de dados do *BCI Competition IV* (2008), Mirnaziri *et al.* (2013) apresentaram um trabalho comparando as taxas de acerto das técnicas de classificação *Multi Layer Perceptron* (MLP), *Support Vector Machine* (SVM), *Linear Discriminant Analysis* (LDA), *K-Nearest Neighbor* (KNN) e *Radial Basis Functions* (RBF) utilizando *Common Spatial Pattern* (CSP) na etapa de pré-processamento. Para o estudo, o sinal de EEG foi filtrado nas bandas de frequência μ , β e γ . A maior taxa de acerto apresentada neste trabalho foi 80,2% com os sujeitos S3 e S8, utilizando o método MLP.

Em pesquisa realizada no Portal de Periódicos da CAPES (que aglutina entre outras bases, IEEE, Springer, Elevator e PubMed), nos últimos cinco anos, foi possível localizar quarenta e três trabalhos relacionados ao Periodograma de Welch e BCI. No mesmo período de tempo foi realizada uma nova pesquisa encontrando treze trabalhos relacionados aos fenômenos ERD e ERS avaliados com o Periodograma de Welch. No trabalho desenvolvido por Machado (2012), por exemplo, utilizando este estimador para os classificadores LDA e NB, atingiu-se uma taxa de acerto de $86,0 \pm 6,5\%$ e $84,8 \pm 6,6\%$.

O trabalho apresentado por Carra (2012), apresentou um sistema *BCI* baseado em ritmos sensório-motores para controle de dispositivos, utilizando PCA (*Principal Component Analysis*), ICA (*Independent Component Analysis*) e CSP (*Common Spatial Pattern*) como características extraída do sinal. Os classificadores selecionados foram o LVQ (*Learning Vector Quantization*) e o DSLVQ (*Distance Sensitive Learning Vector Quantization*), com o *BCI Competition II* atingiu uma taxa de acerto máxima de 74,6% em um sistema automático e 88,2% em um sistema manual.

A Tabela 1 apresenta os principais trabalhos descritos nesta seção.

Tabela 1- Publicações selecionadas para a classificação do EEG.

Autores/ano	Banco de dados utilizados	Classificador Empregado	Melhor Taxa de Acerto
Luo et al. (2019)	<i>BCI Competition IV</i>	SVM associado a <i>Time Series Voting</i>	85,0%
Bhattacharyya et al. (2015)	Do Autor	<i>Type-2 fuzzy</i> embasado em <i>multi-class ANFIS</i>	90,9%
Saha et al. (2015)	Do Autor	Neuro-Fuzzy	95,5%
Hsu (2014)	<i>BCI Competition II e IIIb</i>	<i>Adaptative Fuzzy Neural Network (AFNN)</i> , <i>Support Vector Machine (SVM)</i> e <i>Linear Discriminant Analysis (LDA)</i>	82,3%
Wei et al. (2014)	<i>BCI Competition II</i>	<i>Empirical Mode Decomposition (EMD)</i>	78,0%
Machado e Balbinot (2014)	<i>BCI Competition II</i>	<i>Nave Bayes (NB)</i>	70,0%
Lin et al. (2013)	<i>BCI Competition II</i>	<i>Fuzzy C-mean</i>	82,8%
Mirnaziri et al. (2013)	<i>BCI Competition IV</i>	<i>MSJAD e FSDE</i>	77,0%
Machado (2012)	<i>BCI Competition II</i>	LDA e NB	86,0±6,5%
Carra (2012)	<i>BCI Competition II</i>	LVQ e DSLVQ	88,2%

Fonte: Autor.

2.5 ABORDAGEM SIMPLIFICADA SOBRE PRÉ-PROCESSAMENTO E EXTRAÇÃO DE CARACTERÍSTICAS

As técnicas de extração de características têm como objetivo selecionar as informações relevantes dos dados adquiridos, que representam o sinal original, para permitir a distinção entre os sinais cerebrais de interesse daqueles referentes à atividade cerebral de fundo e de forma a permitir a segmentação desses sinais em classes.

Pode-se utilizar a estimação espectral para descrever a distribuição, em frequência, da potência contida em um sinal, considerando-se um conjunto finito de dados. Os espectros podem ser estimados para detecção de sinais que encontram-se em um meio muito ruidoso (SILVA, 2014).

A Densidade Espectral de Potência (PSD) de um processo aleatório estacionário $x(n)$ é dada pela Transformada de Fourier de Tempo Discreto (OPPENHEIM; SCHAFER, 1999) de sua sequência de autocorrelação. Em termos de frequência normalizada, isto é dado pela Equação (11):

$$P_{xx}(w) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} R_{xx}(m), e^{-jwm} \quad (11)$$

onde P_{xx} é a PSD e R_{xx} é a sequência de autocorrelação de $x(n)$. A PSD, segundo Silva (2014), pode ser descrita em função da frequência f em hertz usando a relação $w = \frac{2\pi f}{f_s}$, onde f_s é a frequência de amostragem, conforme apresentado na Equação (12):

$$P_{xx}(w) = \frac{1}{f_s} \sum_{m=-\infty}^{\infty} R_{xx}(m), e^{-j2\pi f/f_s} \quad (12)$$

Considerando a função de autocorrelação de $x(n)$ dada pela Equação (13):

$$R_{xx}(l) = \sum_{m=-\infty}^{\infty} x(n) x(n-l), \quad l = \pm 1, \pm 2 \quad (13)$$

As unidades da PSD são de potência por unidade de frequência, no caso de $P_{xx}(f)$ as unidades são watts/hertz (SILVA, 2014). Um estimador proposto por Welch (1967), consiste na divisão dos dados de série temporal em segmentos, depois, o periodograma modificado de cada segmento é calculado; por fim, realiza-se a média das estimativas; desta forma, define-se a PSD de Welch.

Outras duas métricas matemáticas foram extraídas do sinal e apresentadas ao classificador, *Root Mean Square* (RMS) e Energia, cujas equações estão apresentadas na Equação (14) e na Equação (15) respectivamente:

$$EEG_{RMS} = \sqrt{\frac{1}{N} \sum_{i=1}^N x^2_i} \quad (14)$$

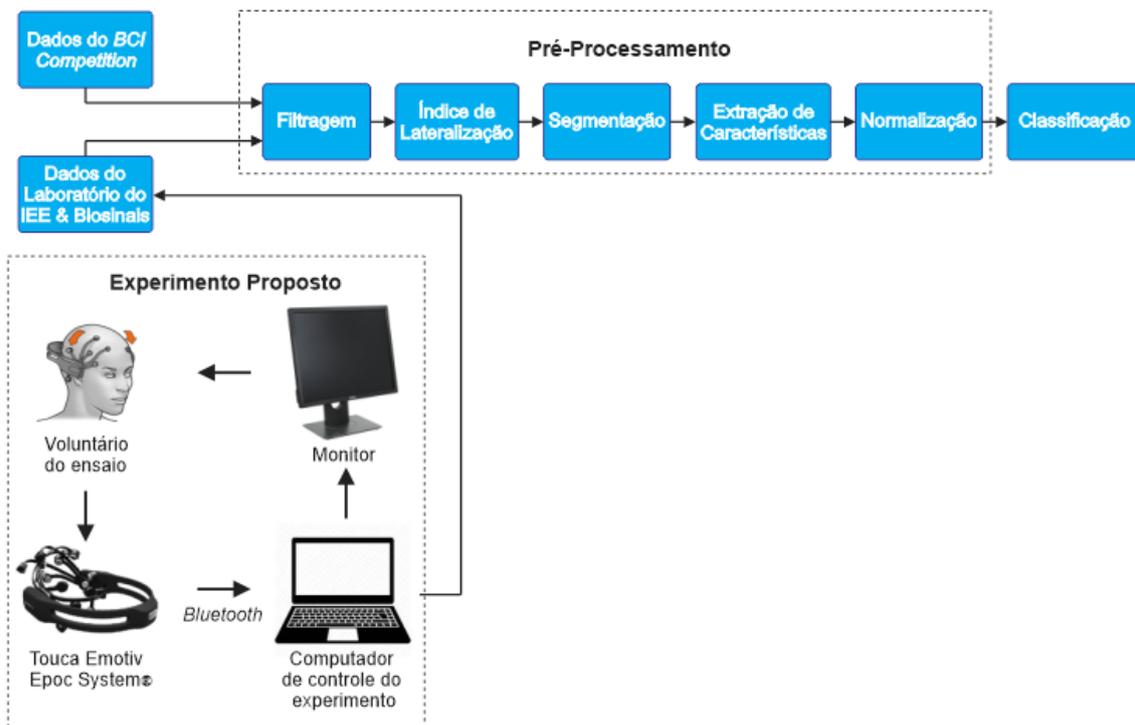
$$EEG_{Energia} = \sum_{i=1}^N |x_i| \quad (15)$$

onde x_i é a janela do sinal onde ocorre a intenção de movimento do indivíduo e N é o número de amostras.

3. METODOLOGIA

Neste capítulo será apresentada a metodologia proposta para a realização deste trabalho. No diagrama de blocos da Figura 7 estão apresentadas todas as etapas propostas neste trabalho. Inicialmente, são utilizados os dados do *BCI Competition* ou um ensaio realizado com o *Emotiv Epoc System®* (MANUAL EPOC, 2019) no laboratório de Instrumentação Eletroeletrônica e Biosinais da UFRGS. Os sinais são salvos em arquivos e, posteriormente, são pré-processados para a extração de características que são apresentadas como entrada dos métodos de aprendizagem de máquina utilizados como classificadores.

Figura 7 – Diagrama de Blocos do sistema proposto.



Fonte: Autor.

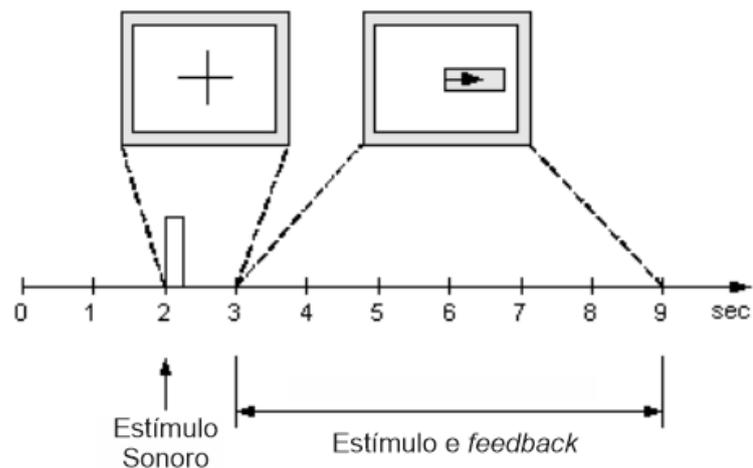
3.1 BANCO DE DADOS DO BCI COMPETITION

Os estudos na área de *BCI* utilizam as mais variadas técnicas de classificação e tipos de características extraídas do sinal, portanto, para efeito de comparação, neste trabalho foram utilizados dois bancos de dados para avaliar o desempenho dos classificadores. Os bancos de

dados foram disponibilizados pela Universidade de Graz, *University of Technology* (PFURTSCHELLER; SILVA, 1999).

O primeiro banco de dados utilizado foi o *BCI Competition II* (2003), composto por duzentas e oitenta trilhas, adquirido a partir de um sujeito saudável do sexo feminino de vinte e cinco anos. Das duzentas e oitenta trilhas, cento e quarenta são de movimentos imaginários para esquerda, e cento e quarenta são de movimentos imaginários para a direita. Para reduzir ruídos e artefatos, os sinais são filtrados com um filtro passa banda na frequência de corte entre 0,5Hz e 30Hz. As trilhas foram adquiridas em sete sessões de quarenta trilhas; cada trilha possui nove segundos, onde no segundo inicial não existe estímulo, aos dois segundos é apresentado um estímulo sonoro e visual (uma cruz) indicando atenção ao voluntário e dos três aos nove segundos é apresentado uma seta para a direita ou para a esquerda como estímulo de movimento imaginário. Na Figura 8 está apresentado como foi executado o procedimento da temporização de cada trilha.

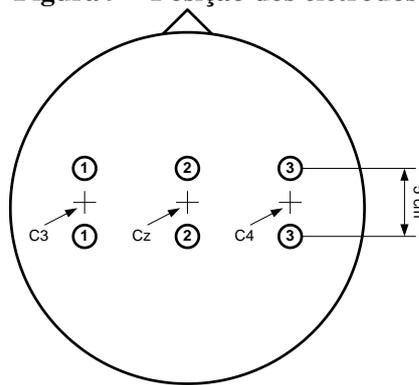
Figura 8 – Temporização das trilhas.



Fonte: Adaptado de *BCI Competition II* (2003).

Os estímulos para o movimento imaginário são apresentados de forma aleatória, os canais bipolares C3, CZ e C4 (ver a correspondente posição na Figura 9) posicionados em uma touca de acordo com o Padrão Internacional do Sistema 10-20, foram utilizados para a correspondente coleta de sinais de EEG. Esses sinais foram adquiridos com uma frequência de aquisição de 128 Hz.

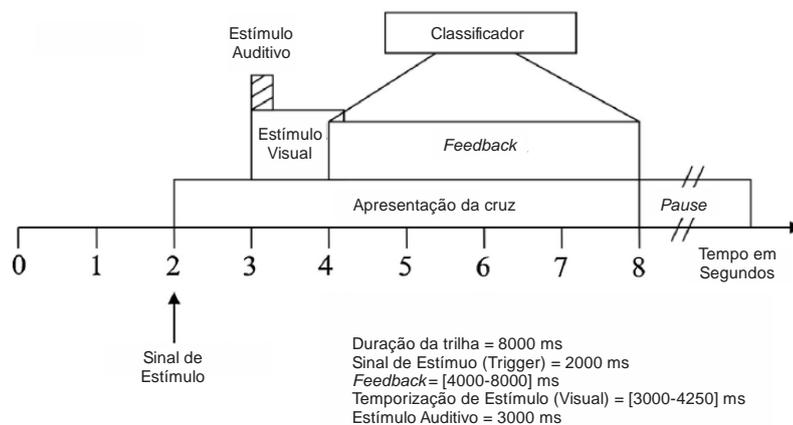
Figura 9 – Posição dos eletrodos.



Fonte: *BCI Competition II* (2003).

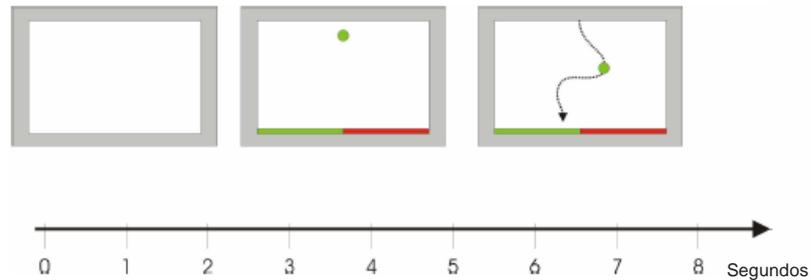
O segundo banco de dados utilizado foi o *BCI Competition IIIb* (2005), composto por três sujeitos (O3, S4 e X11), sendo que a posição dos eletrodos utilizados nesse banco de dados é a mesma da Figura 9, utilizando para esta aquisição um equipamento de EEG com filtro Notch na frequência entre 0,5Hz e 30Hz. A frequência de aquisição desses sinais foi de 125 Hz. O *BCI Competition IIIb* (2005), utiliza dois paradigmas diferentes para os ensaios do experimento. Para o sujeito O3, o paradigma utilizado está apresentado na Figura 10. E, para os sujeitos S4 e X11 foi utilizado o *Basket Paradigm* apresentado na Figura 12. Para o sujeito O3, as trilhas foram adquiridas em dezesseis ensaios de quarenta trilhas totalizando seiscentas e quarenta trilhas e, para os sujeitos S4 e X11, as trilhas foram adquiridas em vinte e sete sessões de quarenta trilhas cada, totalizando mil e oitenta trilhas.

Figura 10 – Temporização das trilhas do sujeito O3 do *BCI Competition IIIb* (2005).



Fonte: Adaptado de *BCI Competition IIIb* (2005).

Figura 11 – *Basket Paradigm*, utilizado no experimento realizado nos sujeitos S4 e X11.

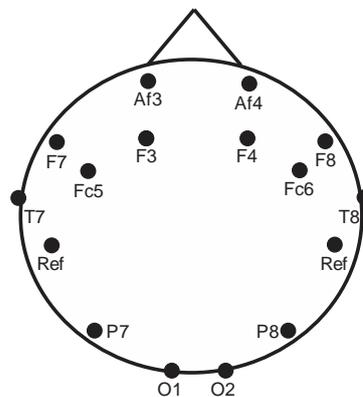


Fonte: Adaptado de *BCI Competition IIIb* (2005).

3.2 BASE DE DADOS GERADA

No Laboratório de Instrumentação Eletroeletrônica e de Biosinais (IEE & Biosinais) da UFRGS, foi realizado um experimento com um voluntário, do sexo masculino, de vinte e três anos de idade, saudável, utilizando um sistema comercial denominado *Emotiv Eloc System*® (MANUAL EPOC, 2019) semelhante ao utilizado por Bhattacharyya *et al.* (2015). Este sistema possui quatorze eletrodos unipolares de superfície que são posicionados no escalpo do voluntário, de forma que neste sistema os eletrodos que se localizam sobre o Córtex Motor do voluntário são os eletrodos FC5 (córtex motor esquerdo), e FC6 (córtex motor direito) conforme apresentado na Figura 12.

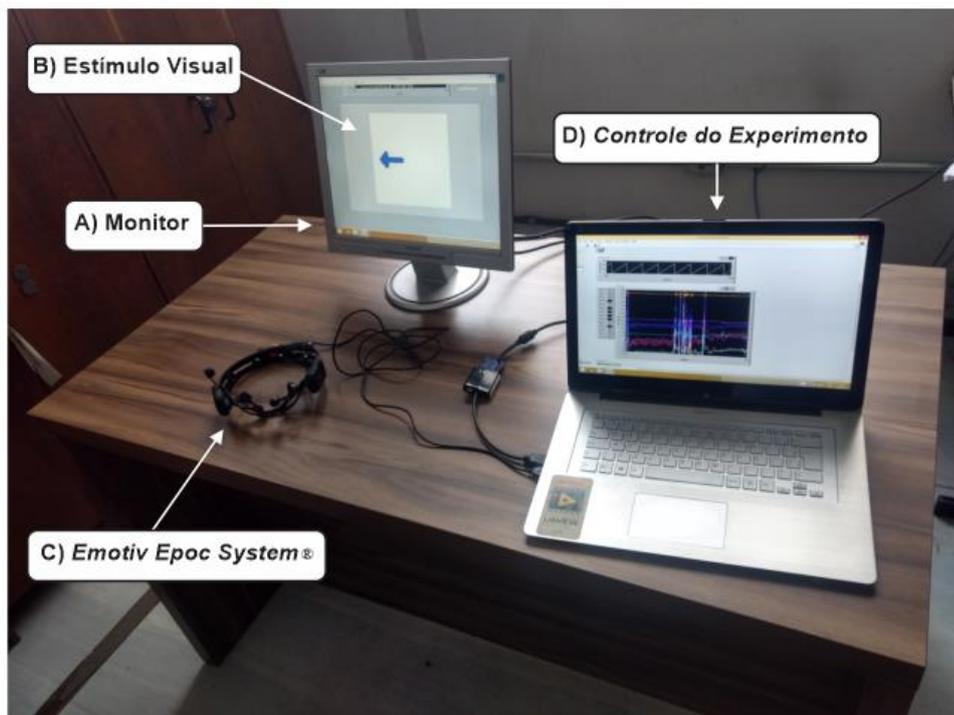
Figura 12 – Posições dos eletrodos no *Emotiv Eloc System*®.



Fonte: Bhattacharyya *et al.* (2015).

As aquisições foram realizadas com um voluntário sentado em uma cadeira com as mãos sobre uma mesa (de forma confortável), nesta mesa encontra-se um monitor de vídeo (Figura 13, item A) onde serão apresentados os estímulos visuais (Figura 13, item B) ao voluntário. O voluntário, utilizando o *Emotive Epoc System*®, nos dois primeiros ensaios deve executar o movimento de flexão da mão e nos seguintes imaginar o movimento de flexão da mão (Figura 13, item C) quando o estímulo visual for apresentado no monitor. O *Emotiv Epoc System*® adquire os sinais de EEG e comunica-se com o computador (Figura 13, item D) via *Bluetooth*. A temporização das trilhas foi feita igual ao experimento do *BCI Competition II* (Figura 8), portanto o protocolo de ensaio é similar (pois o posicionamento dos eletrodos de aquisição difere um pouco) ao experimento realizado pela universidade de Graz. O estímulo sonoro é emitido pelo computador que controla o experimento, conforme apresentado na Figura 13, item D.

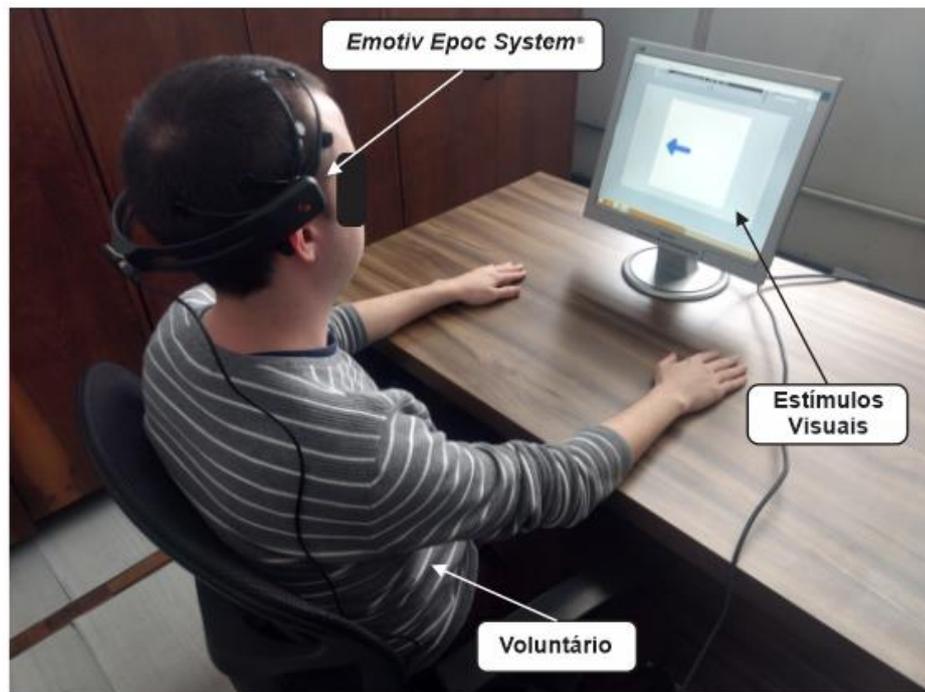
Figura 13 – Aparato utilizado.



Fonte: Autor

A posição do voluntário para os ensaios está apresentada na Figura 15.

Figura 14 – Foto de um dos ensaios.



Fonte: Autor

Os dados são adquiridos pelo *Emotiv Epoc System*® e enviados via *Bluetooth* para o computador. O software *Emotiv Control Panel*® (versão 2.0.0.21) gerencia o recebimento dos sinais de EEG bem como a comunicação *Bluetooth*. As rotinas desenvolvidas no *Labview*® (versão 2014.0.0) comunicam-se com o *Emotiv Control Panel*®, apresentam as telas com os estímulos ao voluntário e salvam os sinais EEG recebidos em arquivos para, posteriormente, serem processados no *MatLab*® (versão 2012). A Tabela 2 resume as informações dos bancos de dados utilizados. É importante observar que, infelizmente, existem diferenças pontuais nos ensaios realizados e cuja influência será analisada na continuidade deste trabalho.

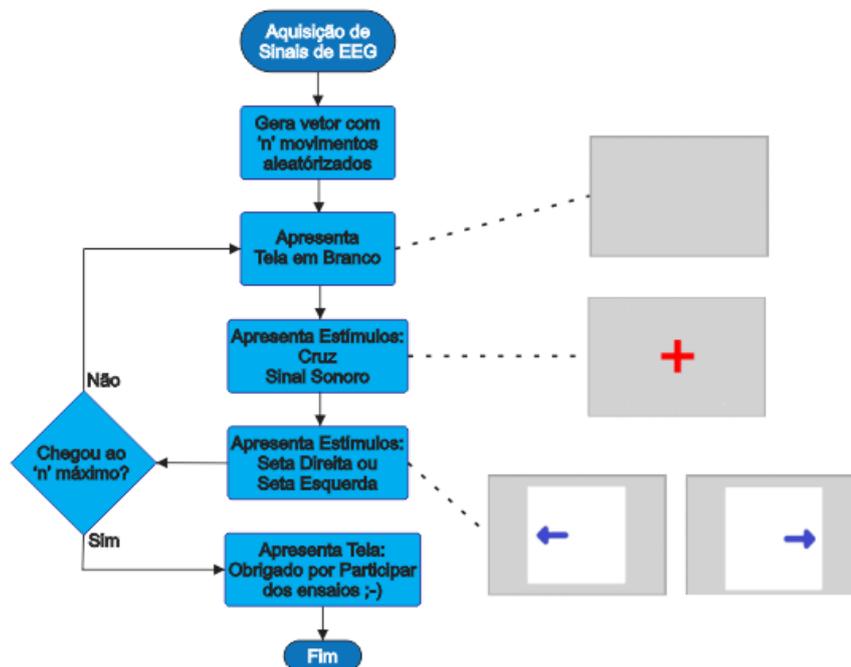
Tabela 2 - Informações dos bancos de dados utilizados.

Dados	Frequência de amostragem	Período de feedback	Tamanho de cada trilha em segundos	Número de voluntários	Número de trilhas
BCI Competition II (2003)	128 Hz	3 a 9s	9s	1	280
BCI Competition IIIb (2005)	125 Hz	4 a 7s	8s	3	640 para O3 e 1080 para S4 e X11
BCI IEE UFRGS (2019)	128 Hz	3 a 9s	9s	1	280

Fonte: Autor.

A aquisição feita a partir da interface *Labview-Emotiv Control Panel*® ocorre com uma taxa de aquisição de 128Hz. A geração do vetor com '*n*' movimentos aleatórios ocorre a partir da criação de dois vetores separados, um com valores um e outro com valores dois, associados aos movimentos de esquerda e direita, respectivamente. Em sequência, os vetores são associados de forma aleatória gerando um vetor de '*n*' posições, divididas igualmente em direita e esquerda. A rotina percorre o vetor de '*n*' posição, associando o número de cada posição com a imagem a ser mostrada ao usuário. Então inicialmente a rotina apresenta uma tela em branco até o segundo número dois, gera um estímulo sonoro de 200 ms e apresenta uma cruz ao voluntário até o segundo três; em sequência é apresentado o estímulo para o movimento imaginário (sendo o número um associado a seta esquerda e o número dois associado a seta direita), até o segundo nove, após isto a rotina reinicia com a tela em branco, e assim sucessivamente, até que tenha sido percorrido todas as posições do vetor '*n*'. O fluxograma da Figura 15 apresenta o procedimento de geração dos estímulos.

Figura 15 – Fluxograma do procedimento de geração do estímulo.



Fonte: Autor

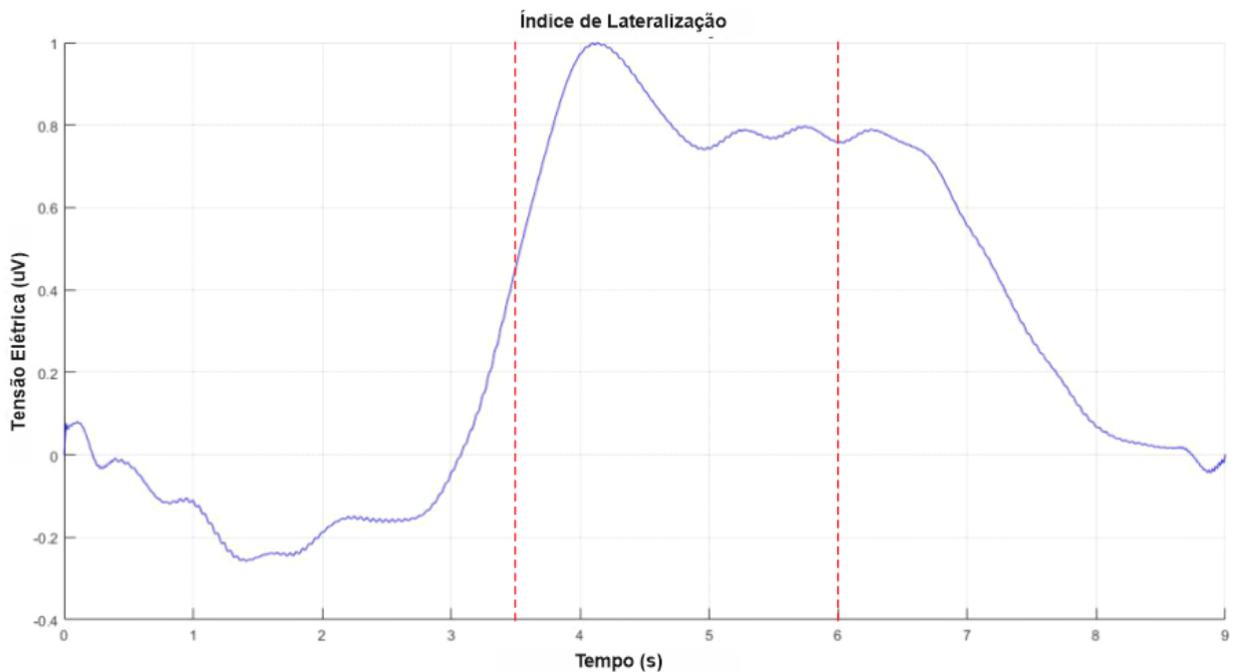
Os sinais de Eletroencefalografia são adquiridos durante toda a execução do software apresentado no fluxograma da Figura 15, e ao final da execução o *Labview*® (versão

2014.0.0) gera dois arquivos: um com os estímulos apresentados ao voluntário junto com o momento no tempo que este estímulo foi apresentado, e o sinal EEG adquirido do voluntário.

3.3 PRÉ-PROCESSAMENTO

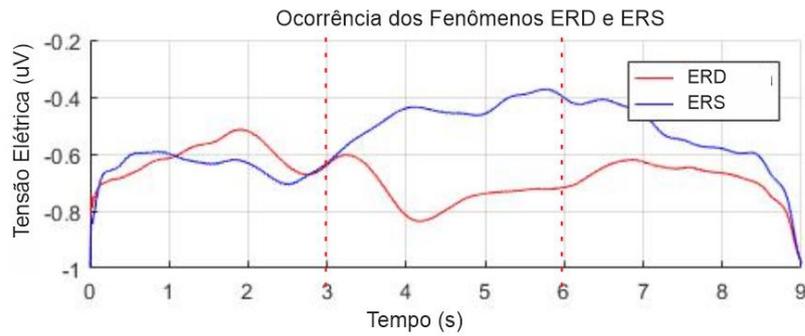
Os sinais dos bancos de dados *BCI Competition* e do ensaio realizado no IEE, foram segmentados conforme o tamanho da trilha de cada um dos seus ensaios. Cada trilha foi filtrada na banda de frequência μ (8 a 11 Hz) e β (26 a 30Hz) com o objetivo de identificar os ritmos sensório-motores e a ocorrência de ERD e ERS. Para filtrar os sinais nessas bandas de frequência, foi utilizado um filtro *Butterworth* de oitava ordem, conforme apresentado por Saha *et al.* (2015). Em sequência é calculado o Índice de Lateralização, com o objetivo de identificar em que período no tempo ocorrem os fenômenos ERD e ERS como ilustrado, por exemplo, na Figura 16. Na Figura 17 está apresentado um exemplo onde pode-se perceber a ocorrência dos fenômenos ERD e ERS e, através da avaliação destes fenômenos, é possível identificar se o indivíduo submetido ao ensaio tem um nível de discriminância suficiente para discernir as classes de interesse (esquerda ou direita).

Figura 16 – Índice de Lateralização calculado a partir do sinal do sujeito S1 do *BCI Competition II*.



Fonte: Autor

Figura 17 – Ocorrência dos fenômenos ERD e ERS a partir do sinal do sujeito S1 do *BCI Competition II*.



Fonte: Autor

3.4 EXTRAÇÃO DE CARACTERÍSTICAS

Portanto, uma vez definido o segmento no tempo onde ocorrem os fenômenos ERD e ERS e, por sua vez, a intenção de movimento do sujeito, o sinal filtrado é segmentado no tempo nessa região. Com base nesse segmento no tempo são extraídas as características do sinal de interesse. As características RMS e Energia foram extraídas conforme a Equação (14) e a Equação (15); o Periodograma de Welch foi obtido dividindo-se a região de interesse nos segmentos mais longos possíveis sem exceder oito segmentos, com sobreposição de 50%. Cada segmento é exibido através de uma janela de Hamming, então, os Periodogramas são calculados para se obter a estimativa do PSD através da Equação (12).

Estas características são normalizadas através da Equação (16) e apresentadas aos classificadores ANFIS e MLP como entradas.

$$X_{normalizado} = \frac{x_i - x_{mínimo}}{x_{máximo} - x_{mínimo}} \quad (16)$$

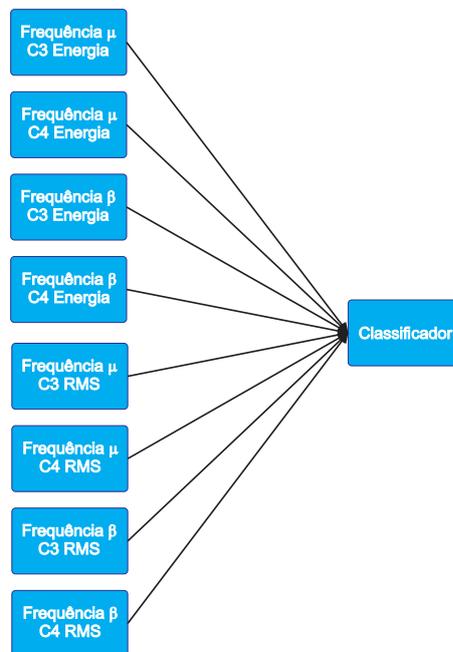
onde $X_{normalizado}$ representa o sinal de interesse normalizado, x_i a janela de interesse a ser normalizada, $x_{máximo}$ o valor máximo do vetor e $x_{mínimo}$ o valor mínimo do vetor.

As características são extraídas do sinal de cada eletrodo (C3, C4 para o *BCI Competition* e FC5, FC6 para o ensaio com a touca *Emotiv EPOC System®*) e as combinações avaliadas foram: Energia; RMS; Periodograma de Welch; Energia e RMS; Energia e Periodograma de Welch; RMS e Periodograma de Welch; RMS, Energia e Periodograma de Welch.

3.5 PROCESSAMENTO E CLASSIFICAÇÃO

Portanto, como as características são extraídas dos sinais filtrados em duas bandas de frequência, o número de entradas de cada classificador varia conforme a combinação de características apresentadas ao mesmo. Quando a característica é apresentada individualmente, o classificador terá quatro entradas, quando houver duas características combinadas, o classificador terá oito entradas e quando houver três características combinadas o classificador terá doze entradas. No diagrama de blocos da Figura 18 está apresentado um exemplo com uma combinação de duas características (considerando os canais C3 e C4 do *BCI Competition*) apresentadas como entrada do classificador.

Figura 18 – Diagrama de blocos de um exemplo de entrada com a combinação de duas características (Energia e RMS).



Fonte: Autor

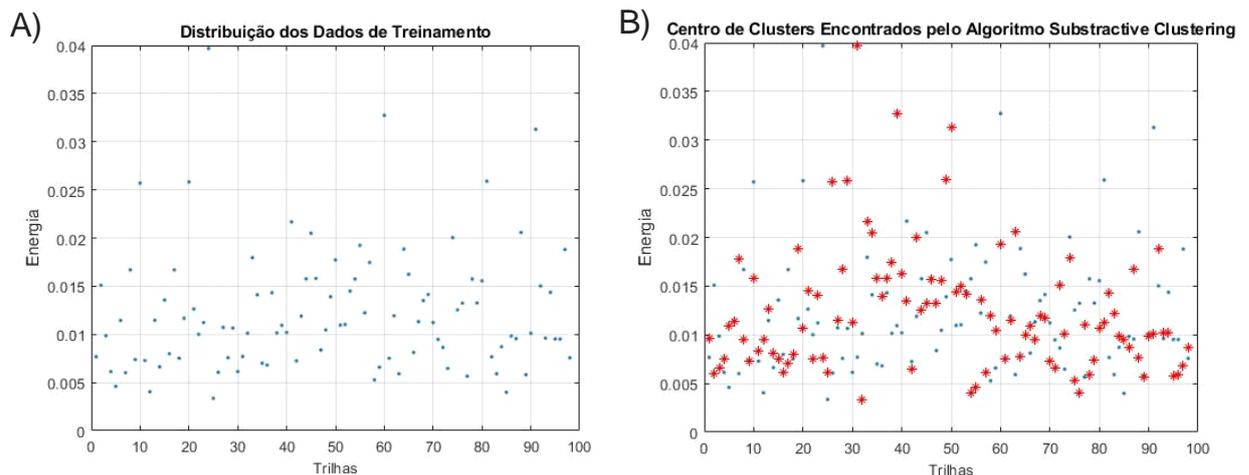
O classificador utilizado com as combinações de características citadas anteriormente foi a *Adaptative Neural Fuzzy Inference System* (ANFIS) com o algoritmo *Subtractive Clustering* (SC), citado no capítulo 2 deste trabalho. Este método foi selecionado devido a utilizar uma técnica de agrupamento de dados, assim como as técnicas apresentadas por Mirnaziri *et al.* (2013) e Bhattacharyya *et al.* (2015), que utilizam *clusters* para auxiliar na classificação dos dados.

A ANFIS não efetua a separação dos dados de treinamento, teste e validação automaticamente, portanto, foi desenvolvida uma rotina na qual as trilhas são separadas

aleatoriamente em 70% dos dados para treinamento e 30% dos dados para teste. Após o teste, as taxas de acerto são calculadas e armazenadas em uma matriz.

O algoritmo *Subtractive Clustering*, com base em uma amostra de dados de entrada e considerando os parâmetros de raio de aceite e raio de rejeição, calcula a densidade dos possíveis centros de *clusters* e seleciona como primeiro aquele que possuir a maior densidade e os demais sucessivamente. Então, considerando um agrupamento de dados de Energia, por exemplo, (ver a Figura 19 A) o algoritmo *Subtractive Clustering* encontra os centros de *cluster* conforme apresentado na Figura 19 B.

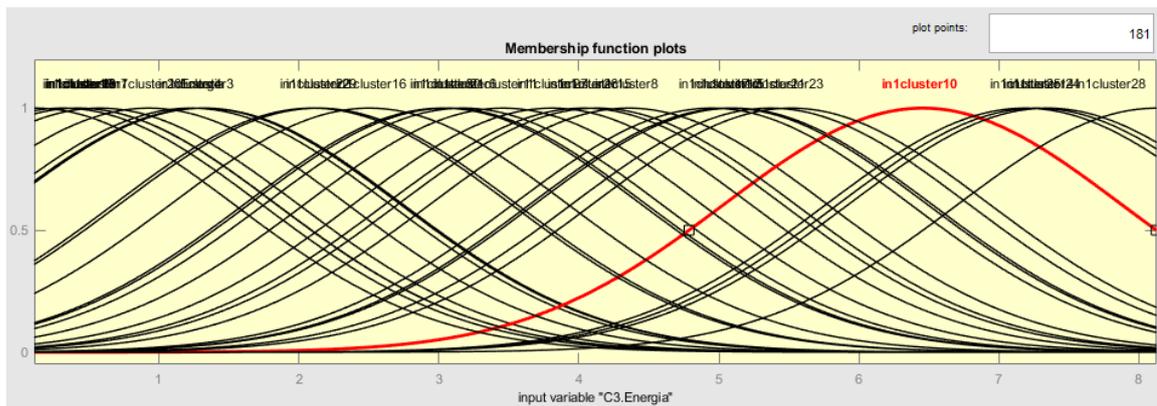
Figura 19 – A) Agrupamento de dados de entrada; B) Centros de *cluster* identificados pelo algoritmo *Subtractive Clustering*.



Fonte: Autor

Na sequência, cada centro de *cluster* encontrado terá uma função de pertinência e uma regra associada ao mesmo. O tipo de Função de Pertinência (FP) selecionada nesse trabalho foi a função gaussiana e os raios dos *clusters* foram selecionados com base em análises preliminares. Por fim, foi selecionado o valor de 0,01 para o raio de aceite e 0,02 para o raio de rejeição. A Figura 20 apresenta as funções de pertinência definidas através do algoritmo *Subtractive Clustering* considerando os dados de Energia apresentados anteriormente.

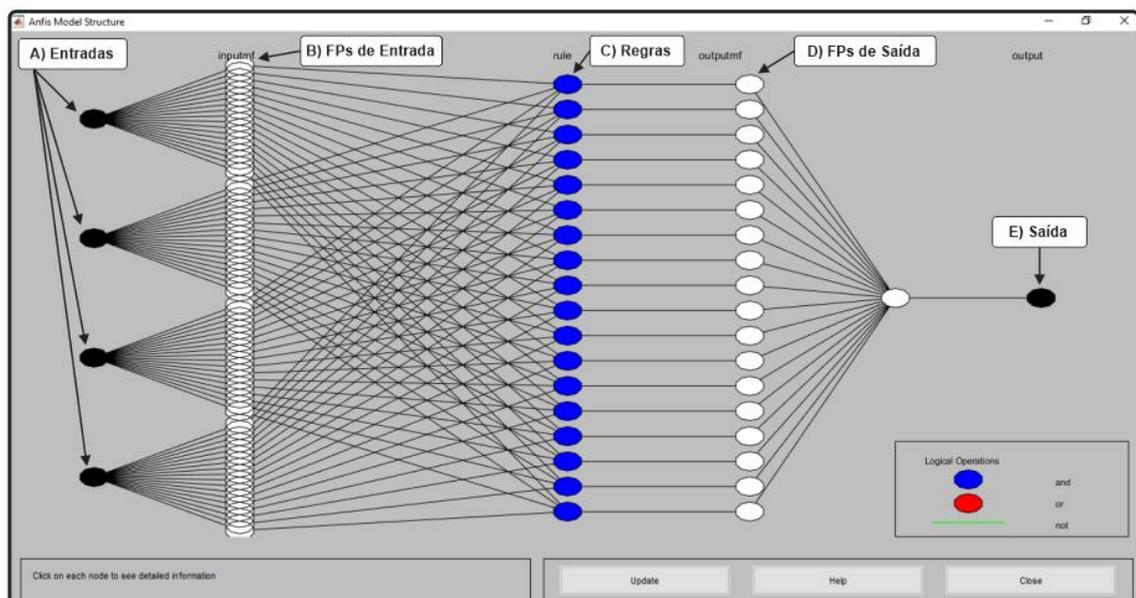
Figura 20 – Funções de Pertinência geradas pelo algoritmo *Subtractive Clustering*.



Fonte: Autor

Como exemplo, a estrutura do exemplo anterior (Figura 19 e Figura 20), elaborada de forma automática pela ANFIS SC, resultou na topologia apresentada pela Figura 21. Nesta Figura 21, é possível destacar as quatro entradas (Figura 21, item A), dezoito Funções de Pertinência para cada entrada (Figura 21, item B), dezoito regras do tipo *And* (Figura 21, item C), dezoito funções de pertinência de saída (Figura 21, item D) e uma saída (Figura 21, item E).

Figura 21 – Exemplo de topologia ANFIS gerada pela técnica *Subtractive Clustering*.



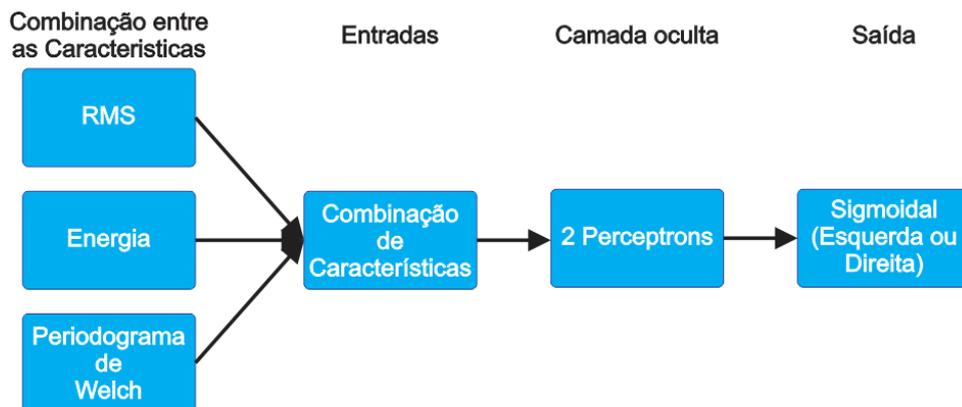
Fonte: Autor

3.6 MÉTODO DE COMPARAÇÃO E ANÁLISE ESTATÍSTICA

Para comparação dos resultados gerados pela ANFIS SC, foi utilizado uma rede neural MLP com o algoritmo de treinamento supervisionado Levenberg-Marquardt. Esta rede neural

foi configurada com quatro, oito ou doze entradas, dependendo da combinação de características selecionadas, semelhante à configuração das entradas da ANFIS SC (a diferença é devido à rede neural MLP possuir um parâmetro configurável de dados para treinamento, teste e validação), dois neurônios na camada oculta (cujas funções de ativação de saída foram do tipo sigmoidal) e uma saída com a função de ativação sigmoidal, conforme apresentado no diagrama de blocos da Figura 22. Os dados de entrada foram configurados com 70% destes para treinamento, 15% para teste e 15% para validação; após a realização dos procedimentos de teste e de validação as taxas de acerto são salvas em um arquivo para posterior análise. O valor selecionado para o parâmetro regularizador λ foi 0,0001 e foi definido com base em testes preliminares. Maiores detalhes sobre a rede neural MLP e o algoritmo de treinamento Levenberg-Marquardt, podem ser consultados no Apêndice A.

Figura 22 – Diagrama de blocos da rede neural proposta.



Fonte: Autor

3.6.1 ANÁLISE ESTATÍSTICA: DOE

Com o objetivo de validar os ensaios estatisticamente, foi realizado um Projeto de Experimentos do tipo Fatorial Completo, tendo como variável de resposta a taxa de acerto e como fatores controláveis a variação de características, os sujeitos, as classes e os métodos propostos neste trabalho, considerando-se que a base de dados do *BCI Competition* segue as premissas de um Projeto Fatorial Completo. O modelo de Projetos de Experimentos segue a Equação (17) e a Tabela 3 mostra os diferentes níveis utilizados para cada um dos fatores controlados.

$$\begin{aligned}
 y_{ijklm} = & \mu + \alpha_i + \beta_j + \gamma_k + \tau_l + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\tau)_{il} + \dots \\
 & \dots + (\beta\gamma)_{jk} + (\beta\tau)_{jl} + (\gamma\tau)_{kl} + (\alpha\beta\gamma)_{ijk} + (\alpha\beta\tau)_{ijl} + \dots \\
 & \dots + (\alpha\gamma\tau)_{ikl} + (\beta\gamma\tau)_{jkl} + (\alpha\beta\gamma\tau)_{ijkl} + \epsilon_{ijklm}
 \end{aligned}
 \tag{17}$$

onde, y_{ijklm} corresponde a m -ésima observação dos níveis i, j, k, l dos 4 fatores controláveis avaliados; $\alpha, \beta, \gamma, \tau$ são os efeitos dos fatores controláveis; μ é a média global da variável de resposta, e ϵ é o erro aleatório que engloba todas as fontes de variabilidade do experimento.

O número de dados analisados é dado pela Equação (18):

$$\begin{aligned}
 Total_{\text{dados}} &= \text{Métodos} * \text{Classes} * \text{Sujeitos} * \text{Características} * \text{Repetições} \\
 Total_{\text{dados}} &= 2 * 2 * 4 * 7 * 30 = 3360
 \end{aligned}
 \tag{18}$$

Tabela 3 – Projeto de Experimentos realizado, baseado nos resultados de classificação.

Taxa de acerto %									
Método ($i = 2$ níveis)	Classe ($j = 2$ níveis)	Sujeito ($k = 4$ níveis)							Totais ($T_{i\dots}, T_{j\dots}, T_{\dots k}$ e $T_{\dots m}$)
		1		...		4			
		Características ($m = 7$)							
		1	...	7	...	1	...	7	
1	1	X_{1111}		X_{1117}		X_{1111}		X_{1111}	$T_{1\dots}$
	2
2	1	:	:	:	:	:	:	:	:
	2	X_{2211}	...	X_{2217}	...	X_{2241}	...	X_{2247}	T_{\dots}

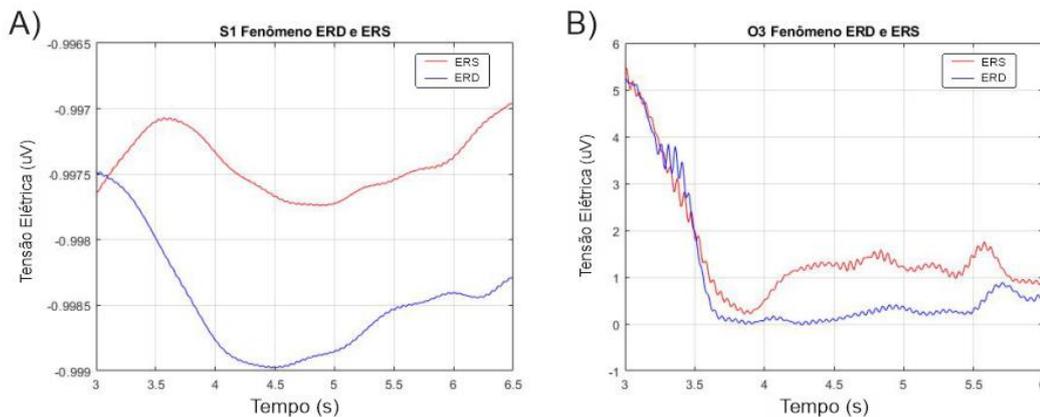
Fonte: Autor

4. RESULTADOS E DISCUSSÕES

Neste capítulo serão apresentados os resultados obtidos utilizando as técnicas descritas nos capítulos anteriores deste trabalho.

A primeira avaliação realizada neste trabalho foi a ocorrência dos fenômenos ERD e ERS em todos os sujeitos dos bancos de dados, com o objetivo de identificar em que momento no tempo, o sujeito submetido aos ensaios conseguiria produzir os fenômenos ERD e ERS, com base nos estudos de Lin *et al.* (2014) e Machado (2012), separando os sinais entre esquerda e direita e somando todos os segmentos conforme apresentado por Pfurtscheller e Silva (1999). Foram identificados os fenômenos ERD e ERS para o sujeito do *BCI Competition II* entre 3 e 6,5 segundos conforme apresentado na Figura 23A. Este sujeito foi o que apresentou a melhor discriminância na geração dos fenômenos durante a imaginação do movimento da mão. Devido à diferença nos paradigmas utilizados nos ensaios, as janelas de tempo onde foram encontrados os fenômenos ERD e ERS foram diferentes. Para o sujeito O3 do *BCI Competition IIIb*, o fenômeno foi localizado entre 3,5 e 6 segundos (Figura 23B).

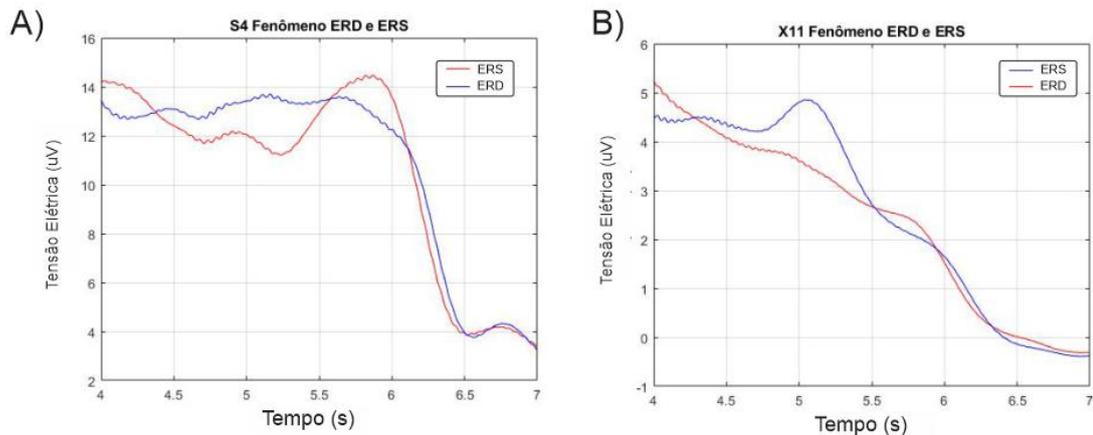
Figura 23 – Ocorrência dos fenômenos ERD e ERS nos sujeitos S1 e O3.



Fonte: Autor

Para o sujeito S4 (Figura 24A) e X11 (Figura 24B) com os quais foi utilizado o *Basket Paradigm*, os fenômenos ERD e ERS foram encontrados entre 4 e 7 segundos.

Figura 24 - Ocorrência dos fenômenos ERD e ERS nos sujeitos S4 e X11.

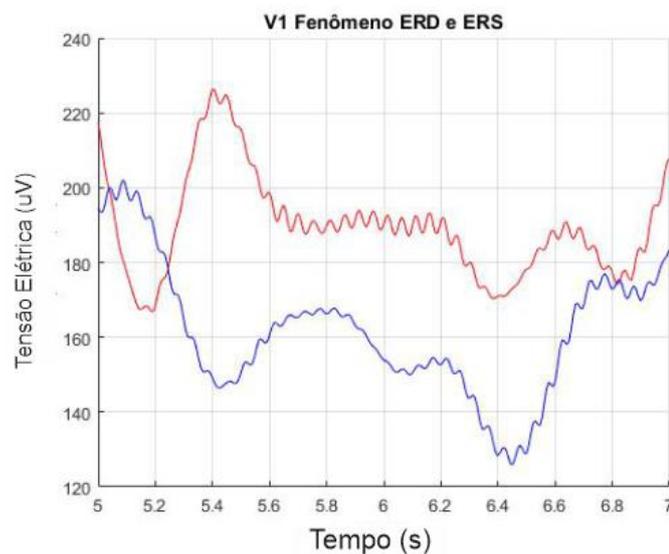


Fonte: Autor

Pode-se observar que a ocorrência dos fenômenos ERD e ERS no sujeito S4 não apresenta-se tão discriminante quanto no sujeito S1. Esta baixa discriminância pode impactar na taxa de acerto.

Para o voluntário V1, cujos sinais foram adquiridos no Laboratório de IEE & Biosinais (UFRGS), os fenômenos ERD e ERS foram encontrados entre 5 e 7 segundos, conforme apresentado na Figura 25.

Figura 25 – Ocorrência dos fenômenos ERD e ERS no sujeito V1 do Laboratório do IEE & Biosinais da UFRGS.



Fonte: Autor

Portanto, em função das diferentes janelas no tempo para discriminação dos eventos ERD e ERA a Tabela 4 resume a segmentação utilizada para cada sujeito avaliado nesse trabalho.

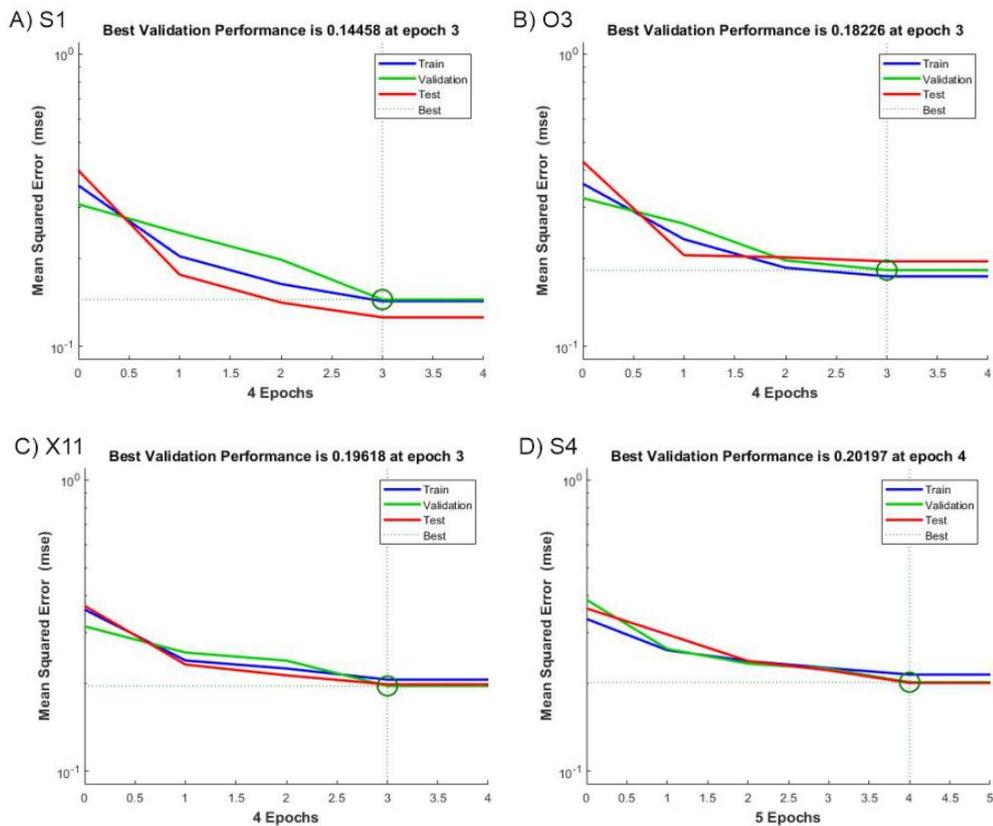
Tabela 4 – Segmentação onde ocorrem os fenômenos ERD e ERS.

Dados	Sujeito	Segmentação no tempo
BCI Competition II (2003)	S1	3,5-6s
BCI Competition IIIb (2005)	O3	3,5-6s
	S4	4-7s
	X11	4-7s
BCI IEE (UFRGS)	V1	5-7s

Fonte: Autor.

Após definida a segmentação onde ocorrem os fenômenos ERD e ERS, desta segmentação foram extraídas as características e treinada a MLP LM com o menor erro quadrático médio como critério de parada. Para os sujeitos S1, O3 e X11 o modelo convergiu e obteve a melhor resposta na terceira época conforme apresentado na Figura 26A, para o sujeito S1; na Figura 26 (item B) para o sujeito O3, e na Figura 26C para o sujeito X11. Para o sujeito S4, o modelo demorou uma época a mais para convergir, conforme apresentado na Figura 26D.

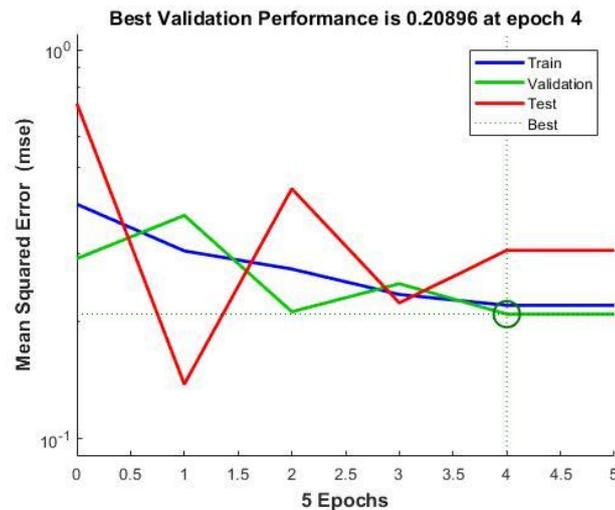
Figura 26 – Performance da RNA MLP LM para os sujeitos do BCI Competition.



Fonte: Autor

O mesmo procedimento foi efetuado com o sujeito V1 do Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS, e embora essa base possua mais ruído que as demais bases avaliadas neste trabalho, a MLP LM conseguiu convergir na quarta época, conforme apresentado na Figura 27.

Figura 27 – Performance do sujeito V1 da base de dados do IEE & Biosinais da (UFRGS).

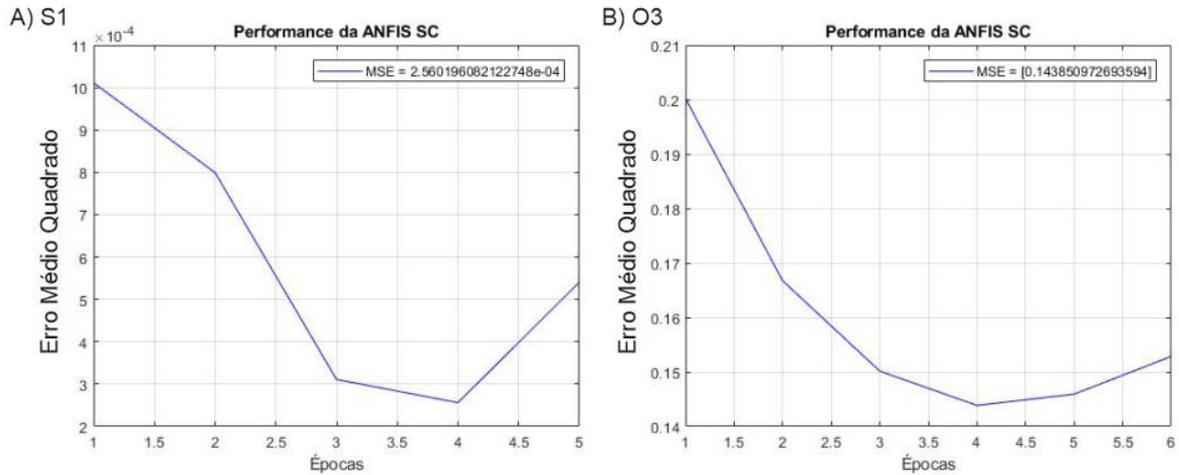


Fonte: Autor

É possível perceber pelos gráficos de performance que o modelo teve uma resposta rápida para todos os sujeitos testados, porém o erro médio quadrático (MSE) do voluntário submetido ao ensaio do Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS ficou acima dos erros dos demais sujeitos, devido a possíveis divergências entre os protocolos executados em cada um dos ensaios como, por exemplo, o posicionamento dos eletrodos, ou o tempo de intervalo entre os ensaios.

Assim como foi feita a segmentação e extração de características para a MLP LM, foi utilizado o mesmo procedimento para a avaliação da ANFIS SC. O critério de parada utilizado também foi o menor MSE. Para o sujeito S1 e para o sujeito O3, o algoritmo SC encontrou cinquenta e quatro *clusters* e gerou cinquenta e quatro regras convergindo na quarta época, conforme apresentado na Figura 28 A e Figura 28 B.

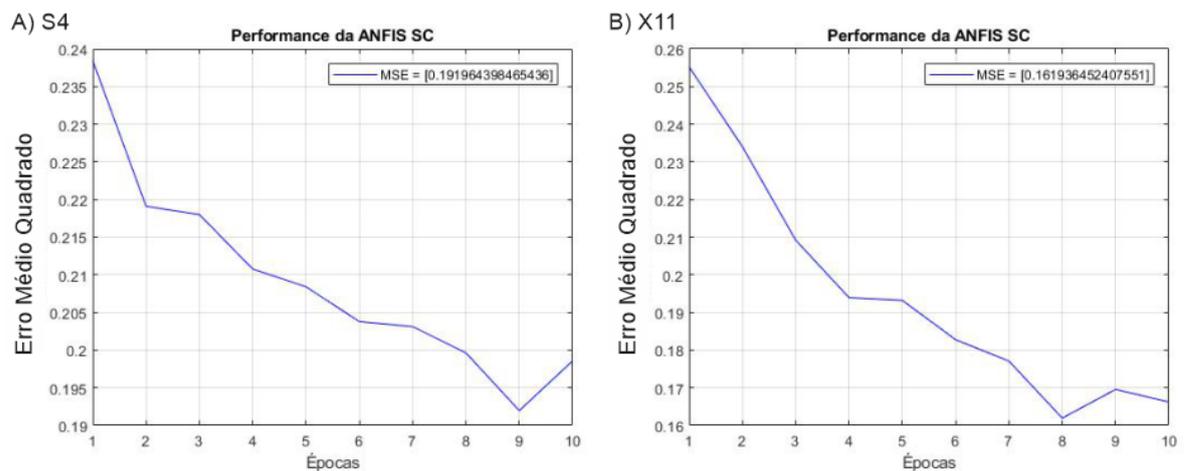
Figura 28 - Performance da ANFIS SC para os sujeitos S1 e O3.



Fonte: Autor

Para o sujeito S4, o algoritmo ANFIS SC encontrou sessenta e oito *clusters* e gerou sessenta e oito regras, chegando ao menor erro médio quadrático na nona época conforme apresentado na Figura 29A, enquanto para o sujeito X11 encontrou sessenta e três *clusters* gerando sessenta e três regras e alcançando o menor erro médio quadrático na oitava época conforme apresentado na Figura 29 B.

Figura 29 - Performance da ANFIS SC para os sujeitos S4 e X11.

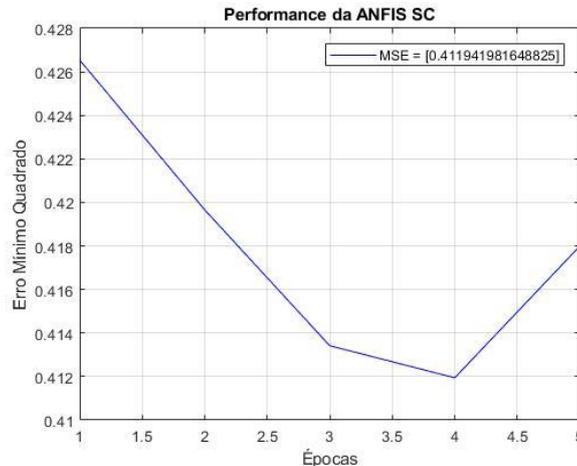


Fonte: Autor

Aplicando o mesmo procedimento para o ensaio do Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS, o algoritmo ANFIS SC encontrou onze *clusters* e

gerou onze regras alcançando o menor erro médio quadrático na quarta época, conforme apresentado na Figura 30.

Figura 30 - Performance da ANFIS SC para os sujeitos V1.



Fonte: Autor

Avaliando a performance da ANFIS SC, a mesma apresentou uma resposta rápida para os sujeitos S1 (*BCI Competition II*) e O3 (*BCI Competition IIIb*) atingindo o menor MSE em quatro épocas e mais lenta para os sujeitos S4 e X11 (*BCI Competition IIIb*). Para o voluntário V1 a rede apresentou uma resposta rápida, mas o menor MSE ficou acima dos erros dos demais sujeitos avaliados.

Ambos os modelos foram rápidos levando em torno de quatro épocas para treinamento, porém para os sujeitos S4 e X11 o modelo ANFIS SC foi mais lento para convergir, no entanto apresentou um valor menor para o mínimo MSE em relação à MLP LM para os mesmos sujeitos. Logo, embora a ANFIS SC seja mais lenta, demonstrou ser mais assertiva para estes sujeitos.

Ainda pode-se determinar uma relação entre o número de *clusters* obtidos no treinamento e a taxa de acerto do classificador (considerando os resultados oriundos dos voluntários do *BCI Competition*). Analisando tais resultados verifica-se que os voluntários que apresentaram menor erro, também foram os que apresentaram o menor número de *clusters* (S1 e O3, os quais obtiveram cinquenta e quatro *clusters*), indicando uma tendência de que quanto maior a discriminância dos sinais, menor o número de *clusters* e, consequentemente, menor o tempo requerido no treinamento do classificador.

Com a rede neural MLP LM configurada, foram realizadas trinta repetições para cada sujeito e com todas as combinações de características possíveis e descritas anteriormente neste trabalho. Os resultados destes testes estão apresentados da Tabela 5 até a Tabela 8.

Nestas tabelas estão apresentadas as taxas de acerto para cada combinação em relação aos sujeitos (com destaque em negrito ao maior valor encontrado).

Tabela 5 - Taxas de acerto para o sujeito S1 utilizando o algoritmo MLP LM.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	84,0±2,7	84,1±5,3	83,8±9,1
RMS	84,8±2,4	81,9±4,8	87,7±5,1
Periodograma de Welch	81,1±7,7	77,2±16	85,1±13
RMS & Energia	83,8±4,5	82,3±7,3	85,3±11
RMS & Periodograma de Welch	84,7±2,7	81,4±5,9	88,0±6,6
Energia & Periodograma de Welch	87,9±1,5	84,4±2,9	91,3±2,8
RMS & Energia & Periodograma de Welch	84,5±2,7	83,5±5,5	85,5±8,2

Fonte: Autor.

Tabela 6 – Taxas de acerto para o sujeito O3 utilizando o algoritmo MLP LM.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	83,5±9,2	83,7±16	83,4±16
RMS	85,7±1,0	85,6±3,7	85,9±4,3
Periodograma de Welch	80,1±12	79,6±23	80,6±23
RMS & Energia	85,7±0,9	85,3±2,8	86,2±3,8
RMS & Periodograma de Welch	86,4±0,6	86,8±4,4	85,9±4,8
Energia & Periodograma de Welch	86,9±0,9	87,5±3,5	86,4±4,2
RMS & Energia & Periodograma de Welch	83,0±13	85,1±7,1	80,9±20

Fonte: Autor.

Tabela 7 – Taxas de acerto para o sujeito S4 utilizando o algoritmo MLP LM.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	65,2±3,4	62,2±14	68,2±12
RMS	65,5±2,5	62,6±5,9	68,4±5,1
Periodograma de Welch	62,8±3,3	62,4±16	63,2±13
RMS & Energia	66,3±0,8	63,1±5,4	69,6±5,6
RMS & Periodograma de Welch	65,9±1,6	62,6±9,1	69,2±7,6
Energia & Periodograma de Welch	67,1±3,2	62,9±12	71,3±8,2
RMS & Energia & Periodograma de Welch	65,7±3,6	61,1±14	70,3±10

Fonte: Autor.

Tabela 8 - Taxas de acerto para o sujeito X11 utilizando o algoritmo MLP LM.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	70,9±1,2	65,5±6,2	76,4±7,0
RMS	71,3±2,6	63,9±4,9	78,6±4,4
Periodograma de Welch	68,7±1,3	61,7±5,5	75,7±6,9
RMS & Energia	70,9±3,7	67,7±7,6	74,3±14
RMS & Periodograma de Welch	71,7±1,3	64,6±4,0	78,7±4,7
Energia & Periodograma de Welch	72,1±1,9	65,7±4,3	78,2±6,6
RMS & Energia & Periodograma de Welch	71,8±2,5	65,6±3,9	78,1±6,6

Fonte: Autor.

Também com a rede neural ANFIS SC configurada, foram feitas trinta repetições para cada sujeito e com todas as combinações de características possíveis e descritas anteriormente neste trabalho, os resultados destes testes estão apresentados da Tabela 9 até a Fonte: **Autor**.

Tabela 12. Nestas tabelas estão apresentadas as taxas de acerto para cada combinação em relação aos sujeitos (com destaque em negrito ao maior valor encontrado).

Tabela 9 - Taxas de acerto para o sujeito S1 utilizando o algoritmo ANFIS SC.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	61,5±7,5	58,6±11	64,4±10
RMS	72,9±5,8	72,9±10	73,0±9,5
Periodograma de Welch	73,9±5,5	73,5±10	74,4±8,0
RMS & Energia	71,1±6,4	70,9±11	71,3±7,5
RMS & Periodograma de Welch	74,6±5,7	76,9±8,7	72,2±11,2
Energia & Periodograma de Welch	90,1±2,0	90,2±3,2	89,9±2,8
RMS & Energia & Periodograma de Welch	76,4±5,2	76,5±9,7	76,3±11

Fonte: Autor.

Tabela 10 – Taxas de acerto para o sujeito O3 utilizando o algoritmo ANFIS SC.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	80,4±3,2	78,9±4,1	81,9±5,8
RMS	79,8±2,5	79,1±4,9	80,5±4,8
Periodograma de Welch	84,3±3,2	84,0±4,4	84,6±6,2
RMS & Energia	76,6±3,9	74,9±5,5	78,3±6,2
RMS & Periodograma de Welch	75,4±5,1	74,8±6,4	76,0±7,6
Energia & Periodograma de Welch	92,6±1,1	92,0±1,5	93,2±2,3
RMS & Energia & Periodograma de Welch	75,1±4,6	75,2±6,0	75,1±6,7

Fonte: Autor.

Tabela 11 – Taxas de acerto para o sujeito S4 utilizando o algoritmo ANFIS SC.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	56,6±3,1	58,4±5,7	54,8±5,6
RMS	54,5±4,0	55,7±6,3	53,3±5,8
Periodograma de Welch	56,1±3,4	56,1±5,3	56,0±5,9
RMS & Energia	54,4±4,3	55,0±5,6	53,9±7,2
RMS & Periodograma de Welch	54,2±3,9	54,9±6,4	53,6±5,9
Energia & Periodograma de Welch	82,9±1,0	84,5±1,7	81,4±1,9
RMS & Energia & Periodograma de Welch	54,3±4,4	54,9±6,9	53,7±5,6

Fonte: Autor.

Tabela 12 – Taxas de acerto para o sujeito X11 utilizando o algoritmo ANFIS SC.

Combinação de Característica	Taxa de acerto Global	Taxa de Acerto Esquerda	Taxa de Acerto Direita
Energia	61,2±3,4	57,7±4,2	64,8±4,5
RMS	60,1±2,9	56,6±5,5	63,7±5,6
Periodograma de Welch	60,5±3,6	56,8±6,3	64,2±5,9
RMS & Energia	56,2±3,7	53,4±5,3	59,0±5,9
RMS & Periodograma de Welch	56,0±4,0	54,5±4,7	57,7±5,8
Energia & Periodograma de Welch	85,7±1,2	84,7±1,8	86,8±1,7
RMS & Energia & Periodograma de Welch	56,8±3,5	54,5±6,4	59,2±5,7

Fonte: Autor.

Avaliando as taxas de acerto para cada combinação de características e realizando uma comparação entre as mesmas, é possível perceber que, para os classificadores e características considerados neste trabalho, a combinação que apresentou a melhor taxa de acerto foi Energia e Periodograma de Welch. Com o classificador ANFIS e o algoritmo *Subtractive Clustering*, a máxima taxa de acerto ocorreu com o sujeito O3 ($92,6 \pm 1,1\%$). Enquanto para o classificador RNA MLP e o algoritmo de treinamento Levenberg-Marquardt, a maior taxa de acerto foi alcançada com o sujeito S1 ($87,9 \pm 1,5\%$).

Considerando o sujeito O3, por exemplo, a ANFIS SC convergiu em quatro épocas com uma taxa de acerto de 92,6%. Tal resultado é promissor se comparado com o de outros trabalhos utilizando a mesma base de dados. O trabalho apresentado por Hsu *et al.* (2014), por exemplo, utilizando EMD MAV FCM levou 10 épocas para convergir e alcançou uma taxa de acerto de 78,0%. Já o trabalho de Lin *et al.* (2014), utilizando o FCM como classificador, com a mesma segmentação no tempo chegou a uma taxa de acerto de 82,0%. O trabalho proposto por Mohammadpour (2016), utilizando *Decision Stump* (DS) e *K-Nearest Neighbors* (KNN) como classificadores e extraindo as características *Autoregressive Parameters* (AR), *Discrete Wavelet Transform*, *Common Spatial Patterns* e *Power Spectrum Density*, alcançou a máxima taxa de acerto de 73,4% com o método KNN e a característica AR.

A Tabela 13 apresenta um ranking resumido dos trabalhos pesquisados em comparação com os resultados encontrados nesta dissertação.

Tabela 13 – Ranking dos trabalhos pesquisados.

Autor	Banco de dados Utilizado	Taxas de Acerto Máximas
Saha <i>et al.</i> (2015)	Do Autor	95,5%
Badaraco (2019)	<i>BCI Competition II e IIIb</i>	92,6%
Bhattacharyya <i>et al.</i> (2015)	Do Autor	90,9%
Carra (2012)	<i>BCI Competition II</i>	88,2%
Machado (2012)	<i>BCI Competition II</i>	88,0%
Luo <i>et al.</i> (2019)	<i>BCI Competition IV</i>	85,0%
Lin <i>et al.</i> (2013)	<i>BCI Competition II</i>	82,8%
Hsu (2014)	<i>BCI Competition II e IIIb</i>	82,3%
Wei <i>et al.</i> (2014)	<i>BCI Competition II</i>	78,0%
Mirnaziri <i>et al.</i> (2013)	<i>BCI Competition IV</i>	77,0%
Mohammadpour <i>et al.</i> (2016)	<i>BCI Competition II</i>	73,4%

Fonte: Autor.

Uma vez comparadas as técnicas apresentadas neste trabalho com as técnicas apresentadas nos trabalhos pesquisados na área, as mesmas foram avaliadas com o ensaio

adquirido no Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS e as taxas de acerto ficaram abaixo do esperado. No entanto, a combinação de características que apresentou a maior taxa de acerto também foi a Energia com o Periodograma de Welch; esta combinação alcançou uma taxa de acerto de 73,8%, para o modelo MLP LM e 71,4% para o modelo ANFIS SC. Estes resultados estão apresentados na **Tabela 14** e, em destaque, encontra-se a combinação com a maior taxa de acerto. Devido às diferenças comentadas anteriormente e à discrepância nos resultados, este sujeito não será considerado na análise estatística, sendo considerado somente os sujeitos oriundos do banco de dados do *BCI Competition II* e *BCI Competition IIIb*.

Tabela 14 – Taxas de acerto máximas para o ensaio do IEE UFRGS.

Combinação de Características	Voluntário 1	
	MLP LM (Maior Taxa de Acerto)	ANFIS SC (Maior Taxa de Acerto)
Energia	52,8%	52,4%
RMS	51,4%	52,4%
Periodograma de Welch	51,4%	52,4%
Energia & Periodograma de Welch	73,8%	71,4%
RMS & Energia	69,0%	61,9%
RMS & Periodograma de Welch	64,3%	54,8%
RMS & Energia & Periodograma de Welch	52,9%	54,8%

Fonte: Autor.

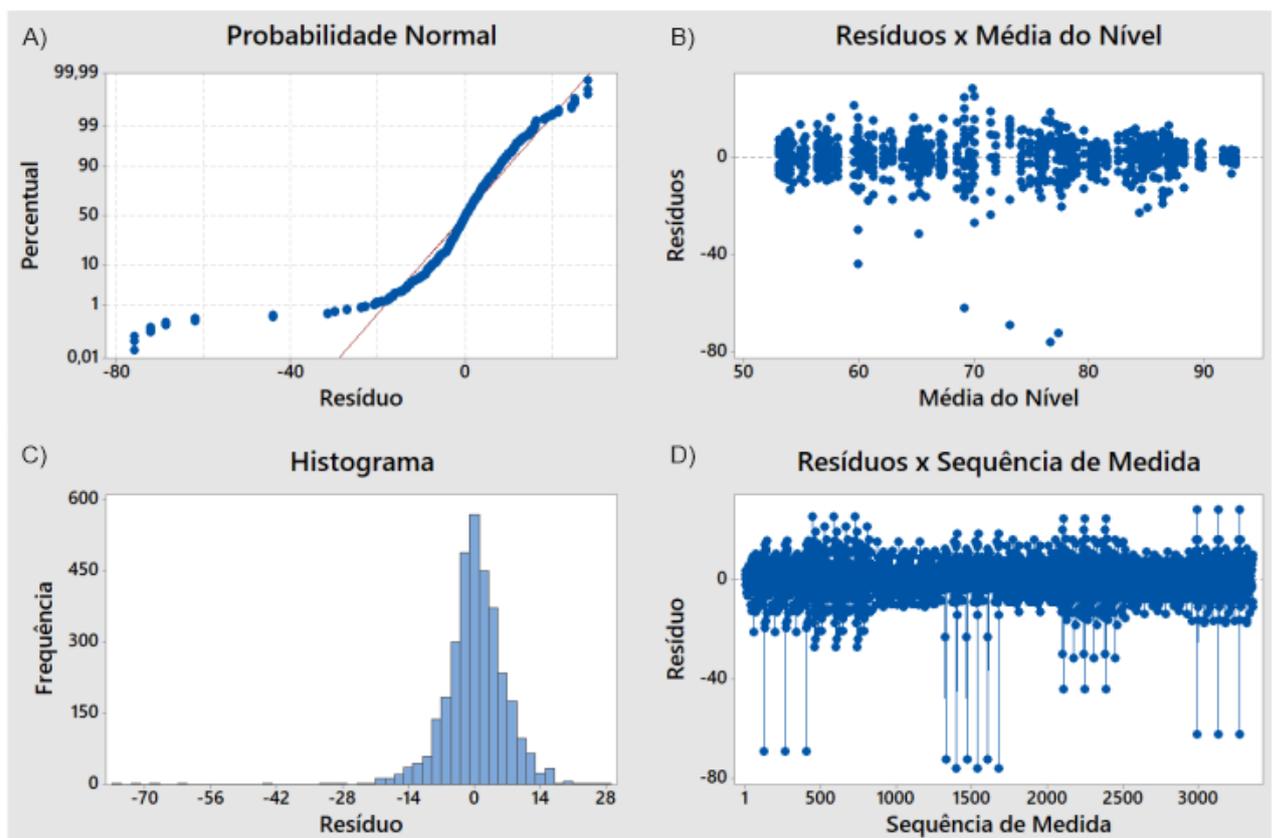
4.1 PROJETO DE EXPERIMENTOS

Com o objetivo de validar o modelo utilizado para analisar o experimento, quatro verificações foram feitas nos dados analisados. A primeira foi embasada na análise do histograma dos resíduos, o mesmo deve seguir uma distribuição normal de média zero. A segunda consiste em verificar a relação dos resíduos em função do tempo (sequência das medidas), para avaliar a existência de correlação entre os resíduos. Nesta análise, a validação ocorre devido à inobservância de correlação, logo a curva citada não deve apresentar nenhum padrão aparente. A terceira está relacionada à análise da dispersão dos resíduos em torno da média dos níveis dos fatores controláveis, que assim como na premissa anterior, não deve apresentar um padrão aparente. A última verificação é a avaliação do gráfico de Probabilidade Normal; este deve apresentar uma tendência linear para o erro estar normalmente distribuído. O correspondente modelo utilizado, assim como, as premissas para sua validação são baseadas em Montgomery (2012).

4.2 VALIDAÇÃO DO PROJETO DE EXPERIMENTOS

Considerando os ensaios com os sujeitos do *BCI Competition*, a Figura 31 apresenta a análise dos resíduos, embora exista a presença de *outliers* o gráfico de Probabilidade normal (Figura 31A) indica uma tendência linear. As relações entre resíduos e a média dos níveis (Figura 31B) não apresentaram uma estrutura aparente, o histograma (Figura 31C), embora deslocado devido à ocorrência de alguns *outliers*, segue uma distribuição normal. Por fim, a relação entre resíduos e sequência de medida (Figura 31D) não apresentou uma estrutura aparente. Com essas observações é possível afirmar que este Projeto de Experimentos tende a ser válido. Em um estudo continuado mais ensaios devem ser realizados para analisar a correspondente Potência Estatística do modelo, assim como uma análise robusta dos *outliers*.

Figura 31 – Análise dos resíduos do Projeto de Experimentos utilizando os dados do *BCI Competition*.



Fonte: Autor

4.2.1 Resultados do Projeto Experimentos

Devido a grande quantidade de interações entre fatores controláveis, selecionou-se, para simplificar a taxa de acerto média obtida como variável de resposta para cada fator controlável. Nesta seção, serão abordados os resultados da influência de cada um dos fatores controláveis (Combinação de Características, Classe, Sujeito e Método), tal qual a interação de todas as combinações entre estes fatores controláveis, considerando a taxa de acerto média obtida. Na Tabela 15 estão apresentados os resultados obtidos do correspondente Projeto de Experimentos com a base do *BCI Competition* com um intervalo de confiança de 95%.

Tabela 15 – Análise de Variância do Projeto de Experimentos do BCI Competition.

Fonte de Variação	Soma dos Quadrados	Graus de Liberdade	Médias Quadradas	Fator F	Fator P	Resultado
Método (Fator A)	38545	1	38544,7	623,46	0,00	Significativo
Sujeito (Fator B)	247661	3	82553,5	7335,31	0,00	Significativo
Características (Fator C)	56222	6	9370,4	151,57	0,00	Significativo
Classe (Fator D)	5716	1	5715,6	92,45	0,00	Significativo
Interação AB	2331	3	777,2	12,57	0,00	Significativo
Interação AC	46036	6	7672,7	124,11	0,00	Significativo
Interação AD	2250	1	2250	36,40	0,00	Significativo
Interação BC	10857	18	603,2	9,76	0,00	Significativo
Interação BD	8544	3	2848	46,07	0,00	Significativo
Interação CD	1045	6	174,2	2,82	0,01	Significativo
Interação ABC	12937	18	718,7	11,63	0,00	Significativo
Interação ABD	3411	3	1137	18,39	0,00	Significativo
Interação ACD	2276	6	379,3	6,13	0,00	Significativo
Interação BCD	4124	18	229,1	3,71	0,00	Significativo
Interação ABCD	5547	18	308,2	4,98	0,00	Significativo
Erro	200803	3248	61,8			
Total	648306	3359				

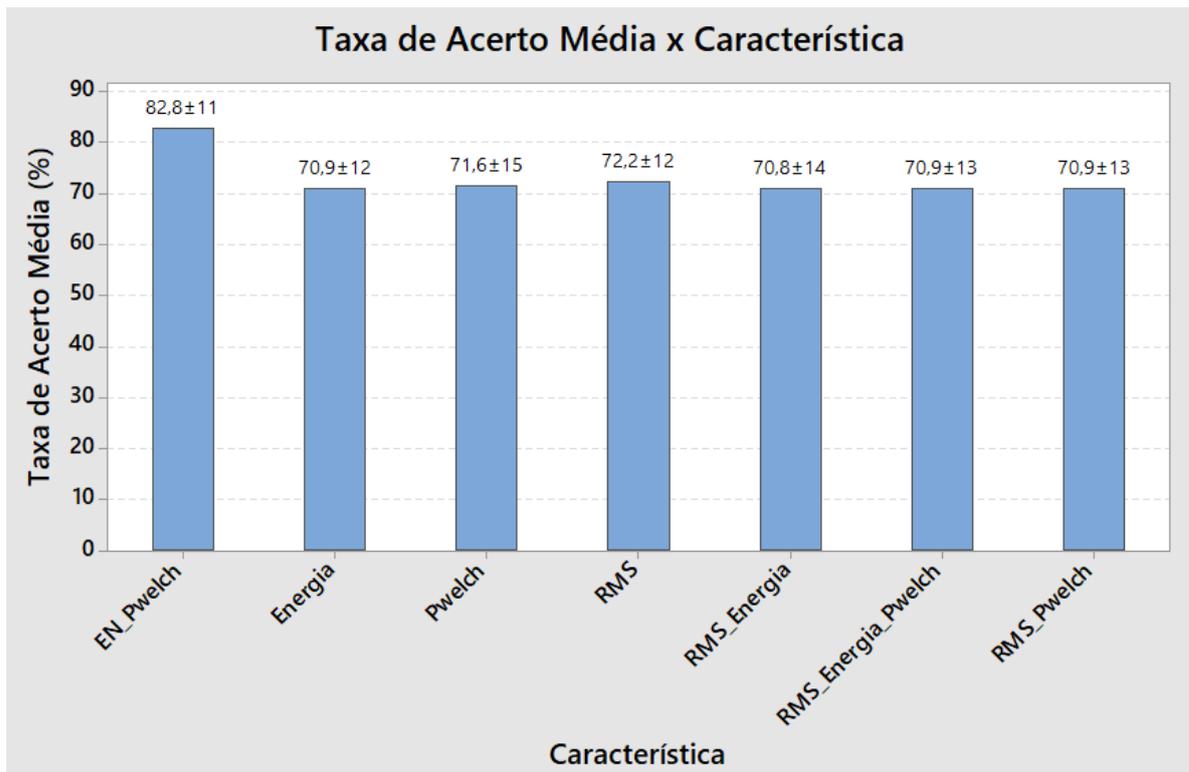
Fonte: Autor.

Observando a Tabela 15, é possível perceber que todos os fatores controláveis e suas interações apresentaram influência significativa na taxa de acerto média, pois em todas as interações, o valor $p \leq 0,5$.

4.2.2 Análise do Fator Controlável Combinação de Características

Na Figura 32 estão apresentadas as médias e o desvio padrão da taxa de acerto de cada combinação de características. Avaliando estes dados, é possível perceber que a combinação de Características: Energia e Periodograma de Welch (EN_Pwelch), foi a que apresentou a maior taxa de acerto, 82,8% e também o menor desvio padrão 11%, considerando os dois métodos e todos os sujeitos. A maior dispersão ao redor da média ocorreu com a característica Periodograma de Welch (Pwelch), cujo desvio padrão resultou em 15%. E a combinação das características RMS e Energia, apresentou a segunda maior dispersão, 14%, e a menor taxa de acerto, 70,9%.

Figura 32 – Análise das Taxas de Acerto das Combinações de Características.



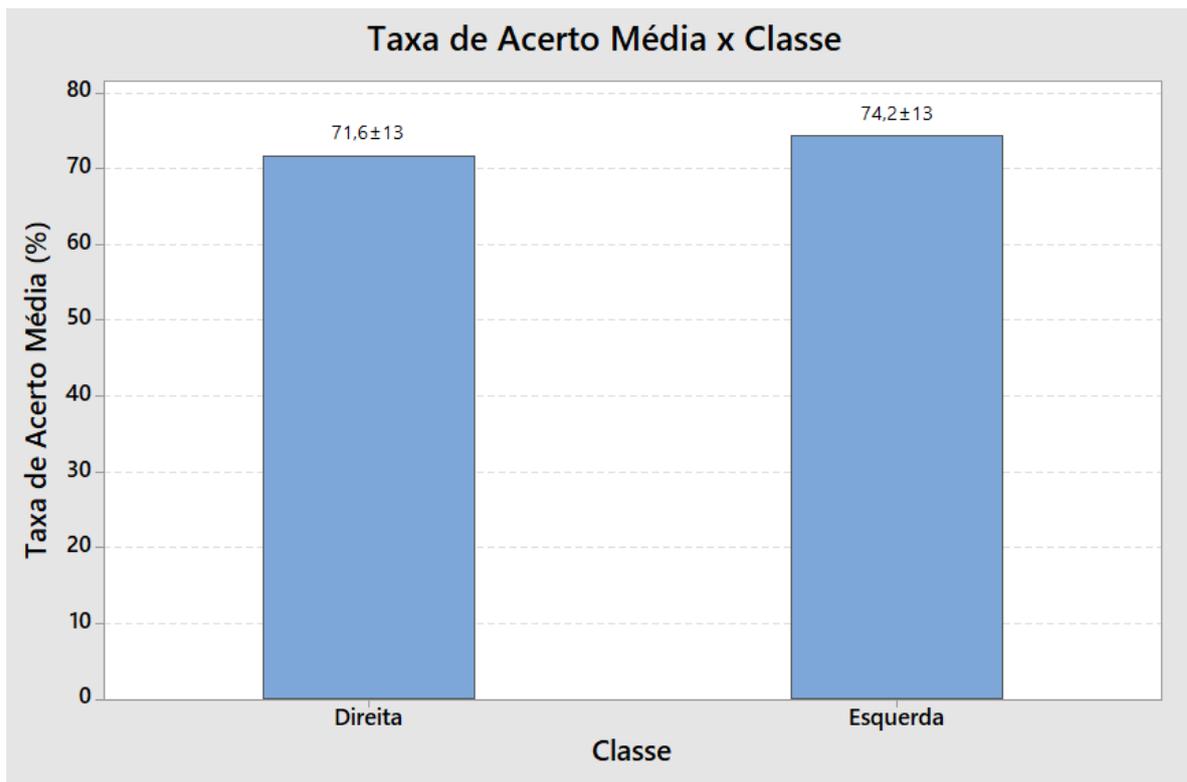
Fonte: Autor

Tais resultados indicam que a combinação de características Energia e Periodograma de Welch, tem a menor variabilidade e a maior assertividade, tornando-se mais promissora com os classificadores e sujeitos avaliados.

4.2.3 Análise do Fator Controlável Classe

As médias e desvio padrão do Fator Controlável Classe estão apresentados na Figura 33. A Classe Esquerda apresentou a maior taxa de acerto, 74,2% e a maior dispersão, 13,9%, enquanto a Classe Direita apresentou uma taxa de acerto de 71,3%, com um desvio padrão de 13,7%. Tais dados foram calculados considerando-se todas as características e métodos, e esta dispersão tão alta ocorreu devido aos ensaios com as características individuais, ou seja, extraindo uma única característica do sinal e apresentando ao classificador. Por exemplo, em um ensaio com o classificador MLP com a característica Energia e o Sujeito O3, o classificador acertou apenas a Classe Esquerda. E, por exemplo, em outro ensaio, com o método ANFIS com o mesmo sujeito e a combinação de características: Energia e Periodograma de Welch, o classificador atingiu uma taxa de acerto de 92,4% para a Classe Direita e 96,2% de taxa de acerto para a Classe Esquerda.

Figura 33 – Análise das Taxas de Acerto do Fator Controlável Classe.

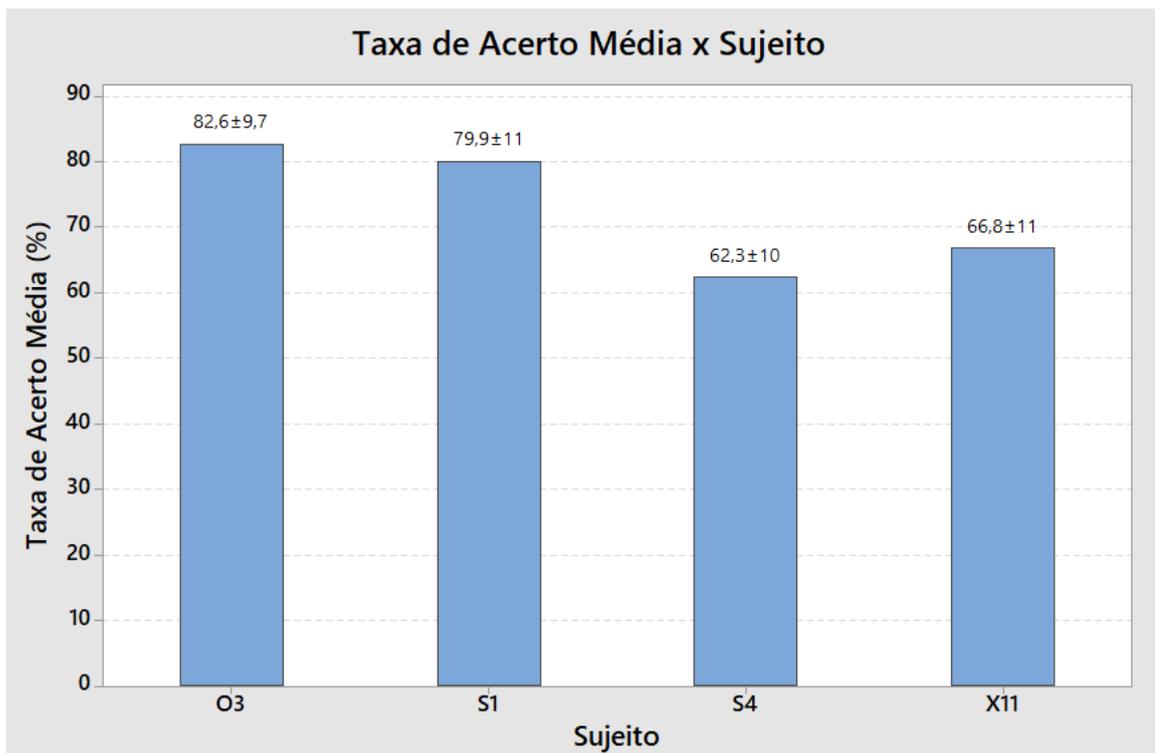


Fonte: Autor

4.2.4 Análise do Fator Controlável Sujeito

Na Figura 34 estão apresentadas as médias e desvios padrões do Fator Controlável Sujeito. O sujeito que apresentou a maior taxa de acerto média e a menor dispersão foi o sujeito O3 do *BCI Competition III* com 82,6% de taxa de acerto média e um desvio padrão de 9,7%, a segunda melhor taxa de acerto foi com o sujeito S1 do *BCI Competition II*, com uma taxa de acerto média de 79,9% e um desvio padrão de 11%. O sujeito que apresentou a menor taxa de acerto foi o S4, pertencente ao *BCI Competition III*, com 62,3% de taxa de acerto e um desvio padrão de 10%. A maior dispersão ocorreu com os sujeitos X11 e S1, com 11%. O sujeito S1 apresentou a segunda menor taxa de acerto com 66,7%.

Figura 34 – Análise das Taxas de Acerto do Fator Controlável Sujeito.



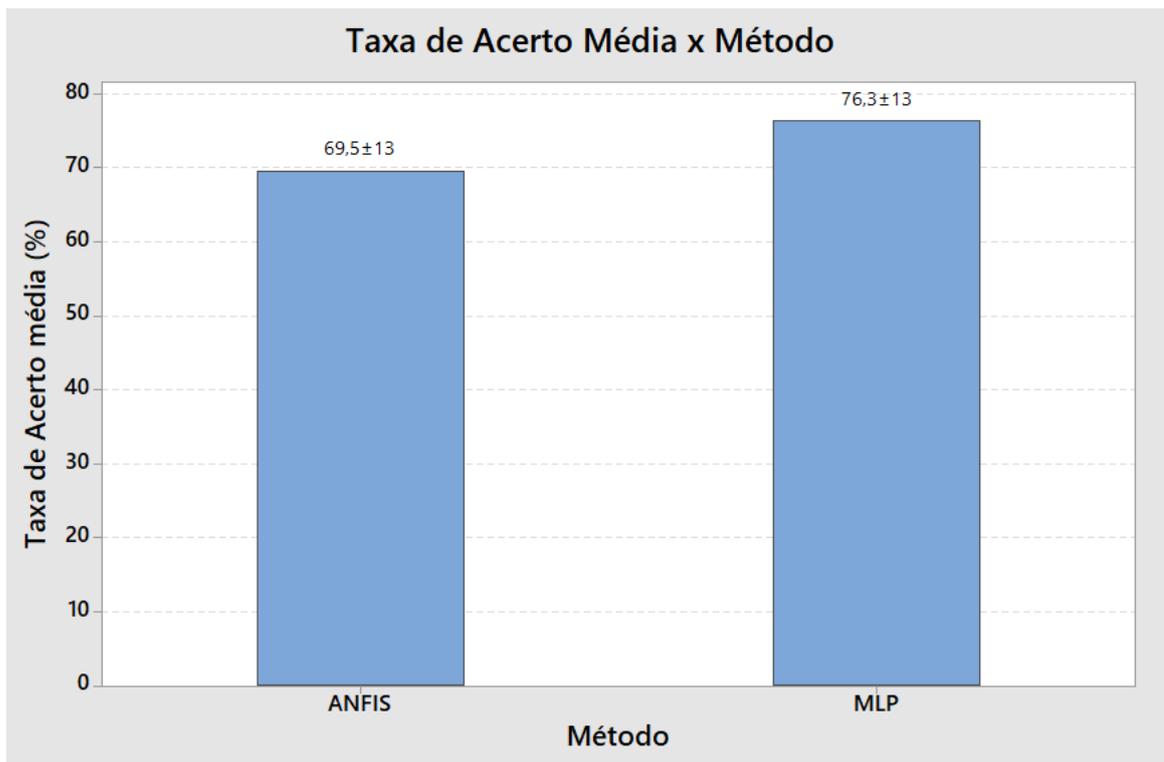
Fonte: Autor

Considerando os gráficos da Figura 23 e da Figura 24, é possível perceber que os sujeitos que apresentaram uma maior capacidade de discriminância ao gerar os fenômenos ERD e ERS, também foram os que, conseqüentemente, apresentaram uma taxa de acerto média maior. Tais resultados são compreensíveis, considerando a capacidade de cada ser humano concentrar-se e, por fim, gerar os sinais de interesse.

4.2.5 Análise do Fator Controlável Método

As taxas de acerto médias e o desvio padrão do Fator Controlável Método estão apresentadas na Figura 35. Estes dados foram calculados considerando todas as características e sujeitos, portanto a dispersão de ambos os métodos ficou alta: 13% para ambos os métodos. O método MLP apresentou uma Taxa de Acerto (76,3%) maior que o método ANFIS (69,5%), pois este método alcançou taxas de acerto com menor variabilidade entre as combinações de características, enquanto o método ANFIS apresentou as melhores taxas de acerto com a combinação de características Energia e Periodograma de Welch. Considerando esta combinação de características, as taxas de acerto do método ANFIS são maiores que as taxas de acerto do método MLP; tais resultados serão detalhados na avaliação da interação entre Método e Características.

Figura 35 – Análise das Taxas de Acerto do Fator Controlável Método.

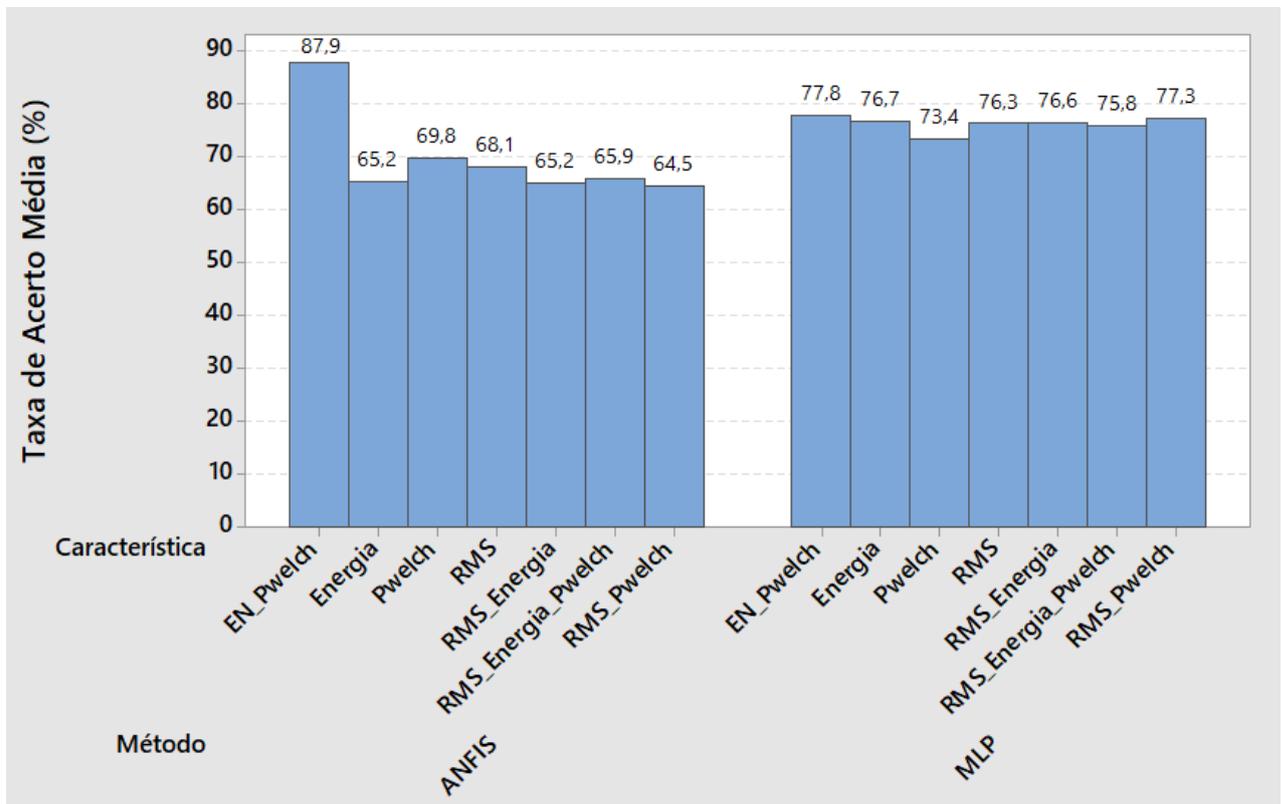


Fonte: Autor

4.2.6 Análise da Interação entre os Fatores Controláveis Método e Características

O gráfico da Figura 36 apresenta a relação entre as taxas de acerto média da interação entre os Fatores Controláveis Método e Combinação de Características. Considerando a combinação de características que apresentou a maior taxa de acerto (EN_Pwelch), o Método ANFIS apresentou a maior taxa de acerto média com 87,9%, enquanto o Método MLP alcançou no máximo 77,8% de taxa de acerto média para a mesma combinação. O método MLP alcançou taxas de acerto maiores que o método ANFIS para as demais combinações. Tais resultados indicam que, para a combinação de características que apresentou a maior assertividade, o Método ANFIS mostrou-se mais promissor que o Método MLP. Para ambos os métodos a combinação das três características juntas (RMS, Energia e Periodograma de Welch) alcançou uma taxa de acerto baixa (65,9% para o Método ANFIS e 75,8% para o método MLP), o que sugere que a combinação de três características para esses Métodos com as configurações propostas dificulta a classificação.

Figura 36 – Interação entre os Fatores Controláveis Características e Método.

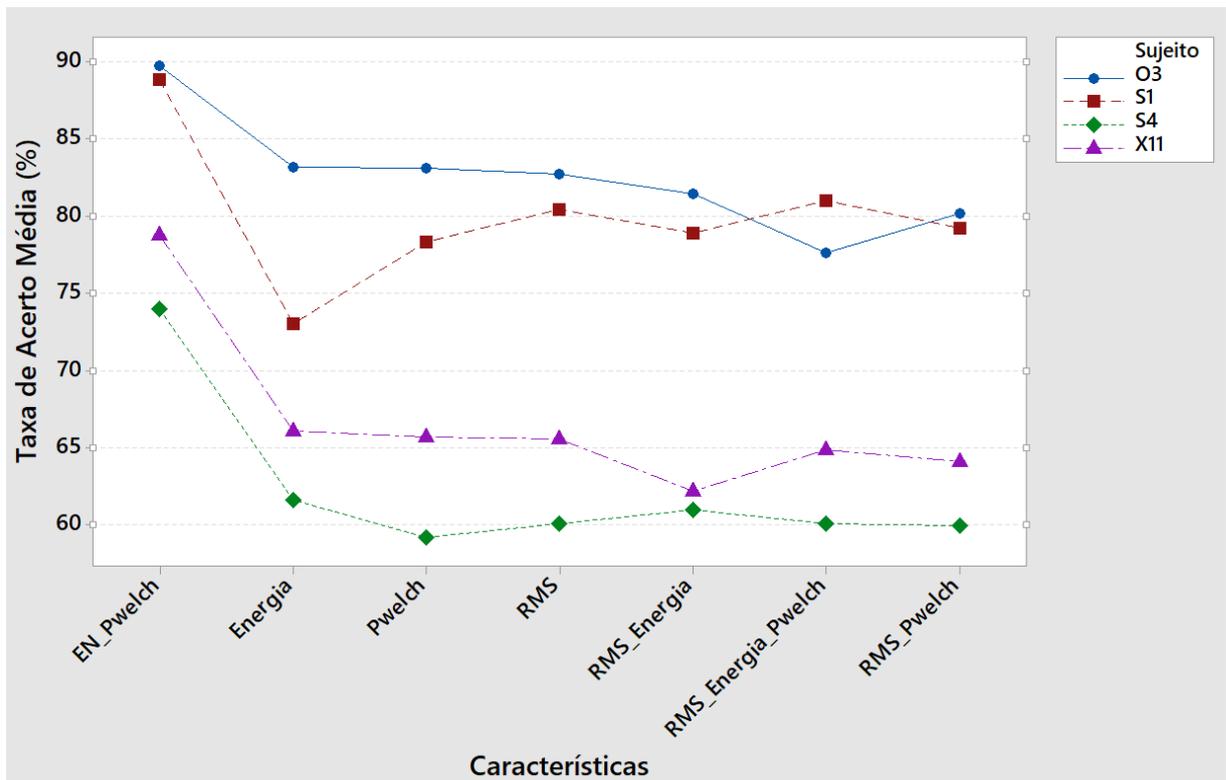


Fonte: Autor

4.2.7 Análise da Interação entre os Fatores Controláveis Sujeito e Características

A Figura 37 apresenta a relação entre os quatro Sujeitos e as Características. Analisando o referido gráfico, observa-se que o sujeito O3 do *BCI Competition III* e o sujeito S1 do *BCI Competition II*, foram os que obtiveram as maiores taxas de acerto em relação aos demais sujeitos, 89,8% e 88,9%, respectivamente, considerando a combinação de características Energia e Periodograma de Welch. Todos os sujeitos obtiveram suas melhores taxas de acerto com esta mesma combinação, sendo que o sujeito X11 atingiu 78,78% e o sujeito S4 alcançou 73,9%. A menor taxa de acerto ocorreu com o sujeito S4 com a característica Periodograma de Welch (Pwelch) alcançando apenas 59,2%. Para o sujeito O3, a menor taxa de acerto ocorreu com a combinação das três características (RMS, Energia e Periodograma de Welch), 77,7% e, para o sujeito S1, a pior taxa de acerto ocorreu com a característica Energia 73,1%.

Figura 37 – Interação entre os Fatores Controláveis Sujeito e Características.



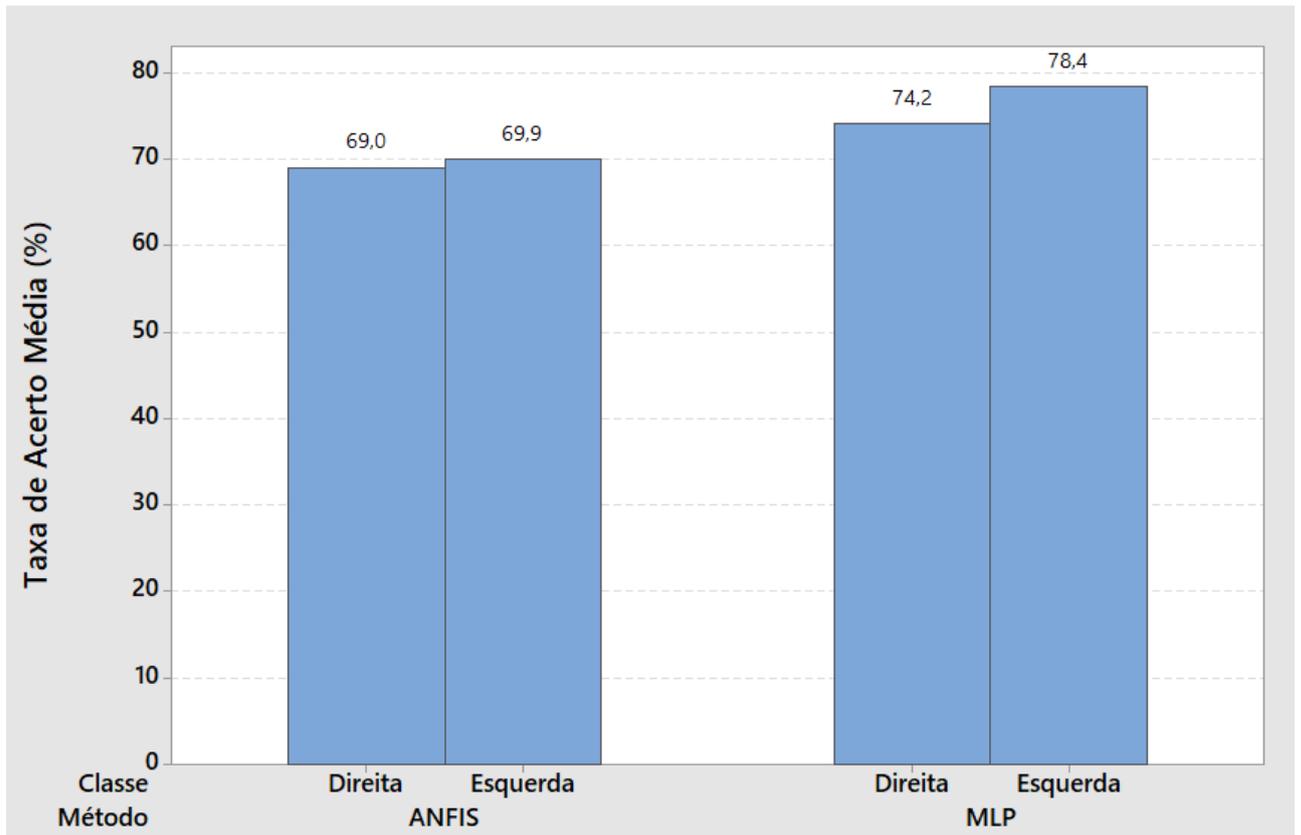
Fonte: Autor

Considerando os dados analisados é possível perceber que a característica extraída do sinal e a capacidade do sujeito gerar os fenômenos ERD e ERS influenciam na taxa de acerto.

4.2.8 Análise da Interação entre os Fatores Controláveis Método e Classe

No gráfico da Figura 38 estão apresentadas as taxas de acerto de cada classe para os dois métodos propostos neste trabalho. Neste gráfico é possível avaliar a interação entre os fatores controláveis Método e Classe, onde é perceptível que ambos os métodos apresentaram assertividade maior na Classe Esquerda, 69,9% para o método ANFIS e 78,4% para o método MLP. O método ANFIS apresentou uma diferença menor entre as taxas de acerto das classes (0,9%) enquanto o método MLP apresentou a diferença de 4,2% entre as taxas de acerto das Classes, o que demonstra que o método ANFIS foi mais equânime em relação à assertividade das classes avaliadas neste trabalho.

Figura 38 – Interação entre os Fatores Controláveis Método e Classe.

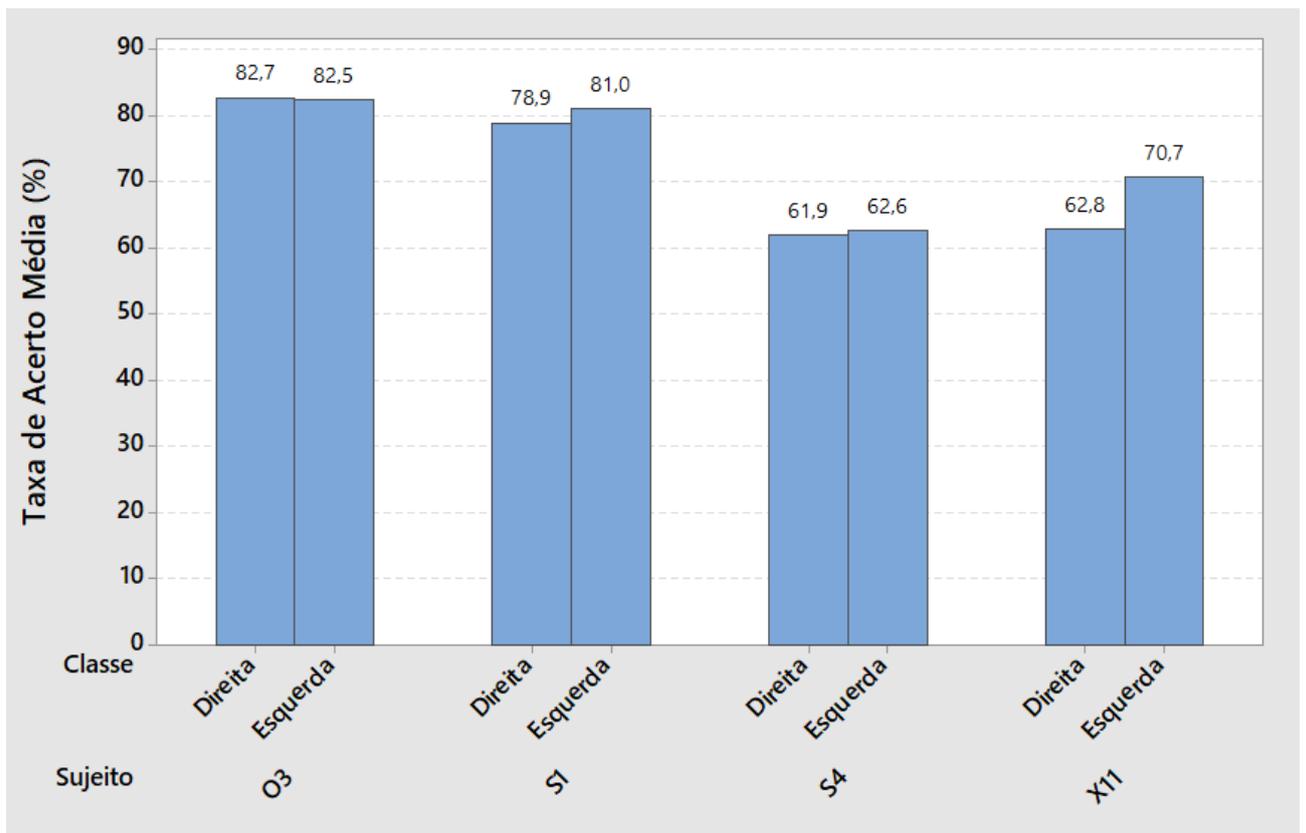


Fonte: Autor

4.2.9 Análise da Interação entre os Fatores Controláveis Sujeito e Classe

Analisando a interação entre Sujeito e Classe, através do gráfico apresentado na Figura 39, é possível perceber que os sujeitos O3 e S4, apresentaram pouca variabilidade nas taxas de acerto entre as classes Direita e Esquerda. O sujeito O3 do *BCI Competition III* apresentou a diferença de 0,26%, sendo 82,7% de acerto para a classe Direita e 82,5% de acerto para a classe Esquerda. E o sujeito S4 apresentou 0,7% de diferença, sendo 61,9% de acerto para a classe Direita e 62,6% de acerto para a classe Esquerda. Os sujeitos S1, S4 e X11 apresentaram maior taxa de acerto para a Classe Esquerda. O Sujeito X11, apresentou a maior diferença entre as taxas de acertos de cada Classe, 7,9% e maior assertividade para a Classe Esquerda, com 70,7%. O sujeito S4 foi o que apresentou a menor taxa de acerto para as duas Classes. Tais resultados reafirmam que a capacidade do sujeito gerar os fenômenos ERD e ERS, influencia as taxas de acerto dos métodos.

Figura 39 – Interação entre os Fatores Controláveis Sujeito e Classe.

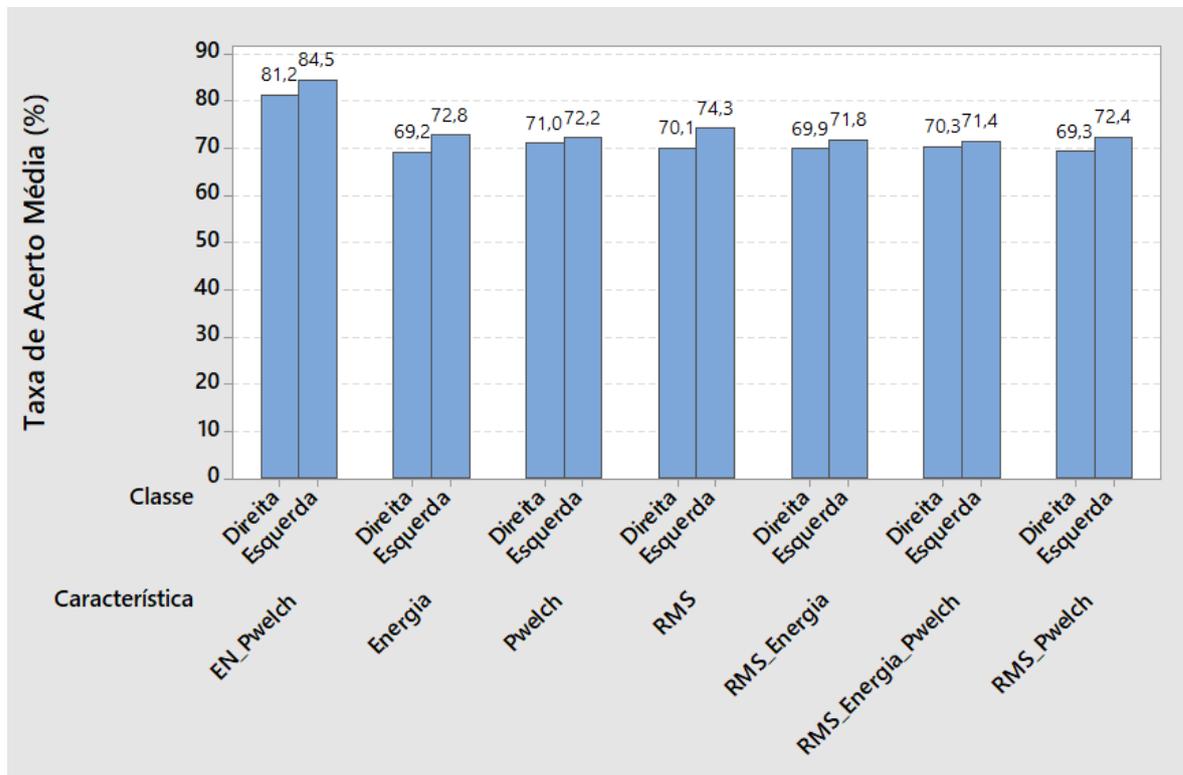


Fonte: Autor

4.2.10 Análise da Interação entre os Fatores Controláveis Característica e Classe

A Figura 40 apresenta a relação entre os Fatores Controláveis Característica e Classe. A partir do referido gráfico, é possível perceber que a interação entre a combinação de características EN_Pwelch (Energia e Periodograma de Welch) e Classe, apresentou o melhor resultado, com a maior taxa de acerto de ambas as Classes e a terceira menor diferença entre estas (1,9%). A menor diferença entre as Classes ocorreu com a interação das três características (RMS, Energia e Periodograma de Welch) com diferença de 1,1%, no entanto esta combinação é a que apresentou a menor taxa de acerto para a Classe Esquerda, 71,4%. A maior diferença entre as taxas de acerto de cada Classe ocorreu com a Característica Energia (3,4%), a qual também apresentou a menor taxa de acerto para a classe Direita, 69,2%. Tais resultados indicam que para a interação Característica e Classe a combinação de características Energia e Periodograma, mostrou-se mais promissora que as demais combinações de características avaliadas neste trabalho.

Figura 40 – Interação entre os Fatores Controláveis Características e Classe.

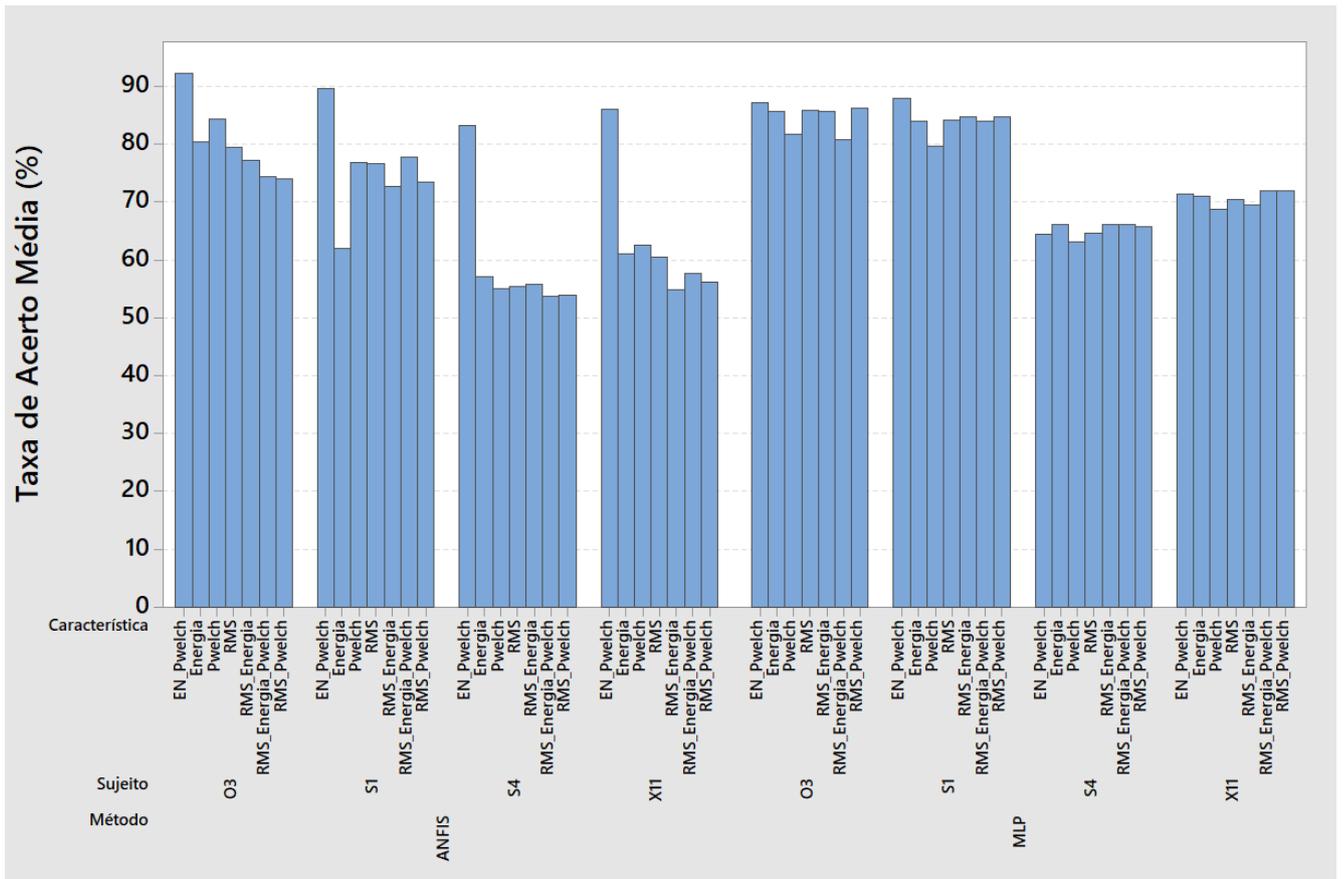


Fonte: Autor

4.2.11 Análise da Interação entre os Fatores Controláveis Método, Sujeito e Característica

O gráfico da Figura 41 apresenta a interação entre os Fatores Controláveis Método, Sujeito e Característica. Neste gráfico está evidenciado que a combinação de características Energia e Periodograma de Welch, obteve a maior taxa de acerto para a maioria dos quatro sujeitos e os dois métodos analisados. Nesta interação apenas o sujeito S4 apresentou a maior taxa de acerto com o Método MLP e a característica Energia, 66,1%. Considerando a combinação de características Energia e Periodograma de Welch, o sujeito O3 foi o que alcançou a maior taxa de acerto, 92,3% para o método ANFIS, enquanto para o método MLP o sujeito S1 foi o que alcançou a maior taxa de acerto, 88,1%.

Figura 41 – Interação entre os Fatores Controláveis Método, Sujeito e Classe.



Fonte: Autor

A menor taxa de acerto ocorreu com o sujeito S1 com o método ANFIS e a combinação das três características juntas (RMS, Energia e Periodograma de Welch) que

atingiu apenas 53,9% de taxa de acerto. Tais resultados evidenciam que a combinação de características Energia e Periodograma de Welch, foi a que apresentou a melhor taxa de acerto. Considerando essa combinação de características, ANFIS foi o método que apresentou o melhor desempenho e considerando todas as características, o método MLP atingiu maiores taxas de acerto para as demais combinações de características. Os gráficos das demais interações estão apresentados no Anexo A.

5. CONCLUSÕES

O propósito principal deste trabalho foi avaliar uma rede *Neuro-Fuzzy* como classificadora de sinais de Eletroencefalografia e comparar suas taxas de acerto com uma rede MLP clássica, utilizando como entrada três diferentes características (RMS, Energia e Periodograma de Welch) extraídas do segmento de interesse do sinal. Estas técnicas de inteligência computacional, foram empregadas como classificadoras do sinal de Eletroencefalografia na identificação de duas classes, esquerda e direita. Estas classes foram criadas a partir da imaginação do movimento de flexão das mãos direita ou esquerda.

Para isso foram utilizadas três bases de dados, duas do *BCI Competition* (*BCI Competition II* e *BCI Competition IIIb*) e um ensaio adquirido no Laboratório de Instrumentação Eletroeletrônica & Biosinais da UFRGS, totalizando cinco voluntários.

Nos trabalhos pesquisados foram encontradas muitas técnicas explorando movimentos imaginários a partir do estudo do sinal no domínio da frequência, no entanto, no presente trabalho foi possível verificar que a combinação de características, uma no domínio do tempo e a outra no domínio da frequência, alcançou taxas de acerto equiparáveis ou melhores que as taxas de acerto de trabalhos na área, conforme apresentado na Tabela 13.

Considerando a análise estatística dos resultados, foi possível validar o Projeto de Experimentos e perceber que a combinação das características Energia e Periodograma de Welch foi a que apresentou o melhor resultado e também que a capacidade de cada sujeito gerar os fenômenos ERD e ERS pode influenciar as taxas de acerto dos modelos avaliados neste trabalho. Dentre os métodos avaliados, considerando a combinação Energia e Periodograma de Welch, o método ANFIS apresentou os melhores resultados e a maior taxa de acerto ($92,6 \pm 1,1\%$ com o sujeito O3). O método MLP apresentou uma diferença menor nas taxas de acerto entre as diferentes combinações de características. Mas a maior taxa de acerto também foi com a combinação de características Energia e Periodograma de Welch ($87,9 \pm 1,5\%$ com o sujeito S1), mas não foi superior à taxa de acerto do método ANFIS com esta mesma combinação de características.

A principal crítica que se faz aqui é quanto ao ensaio efetuado no Laboratório de Instrumentação Eletroeletrônica & Biosinais, devido a possível diferença no intervalo entre os ensaios e diferença entre os tipos e posição dos eletrodos, as taxas de acerto foram menores, o que já era esperado ao avaliar a performance dos classificadores com esse voluntário durante os testes preliminares. O desenvolvimento de um sistema com *feedback* pode aumentar as

taxas de acerto, tendo em vista que nos ensaios onde o voluntário movimentava a mão e não somente imaginava o movimento da mesma, obteve taxas de acerto melhores. Outro fator que pode contribuir com o aumento da performance do classificador é o tempo de intervalo entre ensaios, uma vez que o ensaio que apresentou maior taxa de acerto, foi o que teve um maior tempo de repouso entre ensaios. Na literatura não foram encontrados registros com essa informação.

Do ponto de vista do voluntário foi percebido que a utilização do *Emotive Epos System*®, durante muito tempo gerou desconforto, pois o tamanho da touca e a posição dos eletrodos ao longo do tempo começam a pressionar a cabeça do usuário. Se o sistema for pensado para uma aplicação cotidiana como controle de uma cadeira de rodas ou outro dispositivo qualquer, um sistema que considere o formato e tamanho do crânio do usuário deve ser desenvolvido.

6. PROPOSTA PARA TRABALHOS FUTUROS

Para trabalhos futuros seria importante aumentar o número de voluntários na aquisição de sinais e desenvolver um sistema com *feedback*, que permita ao usuário verificar se está conseguindo gerar os fenômenos ERD e ERS conforme sua intenção de movimento. E que o tempo de repouso entre ensaios seja controlado e seja avaliado o quanto pode impactar na taxa de acerto do modelo. Sugere-se também avaliar as técnicas apresentadas neste trabalho em um banco de dados com mais voluntários como o *BCI Competition IV* (2008), por exemplo, que possui nove sujeitos.

A área de *BCI* ainda é um campo em estudo com diferentes técnicas sendo testadas, como as técnicas apresentadas neste trabalho, uma padronização das temporizações de trilhas e protocolo de ensaio, auxiliaria na comparação entre ensaios, o próprio banco de dados do *BCI Competition IIIb*, por exemplo, apresenta diferença de ensaios em um mesmo banco de dados, portanto estudos que definam uma padronização neste tipo de ensaio serão promissores.

REFERÊNCIAS BIBLIOGRÁFICAS

- ANTELIS, J. M. *et al.* Dendrite morphological neural networks for motor task recognition from electroencephalographic signals. **Biomedical Signal Processing and Control**, v. 44, p.12-24, July 2018.
- BAMDADIAN, A. *et al.* The predictive role of pre-cue EEG rhythms on MI-Based BCI classification performance. **Journal of Neuroscience Methods**, v. 235, p.138-144, June 2014.
- BHATTACHARYYA, S. *et al.* Interval type-2 fuzzy logic based multiclass ANFIS algorithm for real-time EEG based movement control of a robot arm. **Robotics and Autonomous Systems**, n. 68, p. 104-115, 2015.
- BCI COMPETITION III**. 2003. Disponível em: <http://www.bbc.de/competition/ii/> Acesso em: 20 ago. 2018.
- BCI COMPETITION III**. 2005. Disponível em: <http://www.bbc.de/competition/iii/> Acesso em: 20 jun. 2019.
- BRAGA, A. P.; CARVALHO, A. P. L. F.; LUDERMIR, T. B. **Redes neurais artificiais**. 2. ed. Rio de Janeiro: LTC, 2007.
- CARRA, M. **Desenvolvimento de uma interface cérebro computador baseada em ritmos sensório-motores para controle de dispositivos**. 2012. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.
- CHEN, Y.; ZHANG, S. Research on EEG classification with neural networks based on the Levenberg-Marquardt algorithm. New York: Springer, 2012. p.195-202.
- CHIU, S. L. Fuzzy model identification based on cluster estimation. **Journal of Intelligent and Fuzzy Systems**, v. 2, n. 3, p. 267-278, Sept. 1994.
- GARDNER, M. W.; DORLING, S. R. Artificial neural networks (the multilayer perceptron): a review of applications in the atmospheric sciences. **Atmosferica Environment**, v. 32, n. 14-15, p. 2627-2636, Aug. 1998.
- GOBLE, L. **The Blackwell guide to philosophical logic**. Hoboken: Blackwell, 2001.
- GUYTON, A. C.; HALL, J. E. **Tratado de fisiologia médica**. 11. ed. Rio de Janeiro: Elsevier, 2006.
- HAIKIN, S. **Neural networks and learning machines**. 3rd. ed. Ontario: Pearson, 2009.
- HSU, W.-Y. Motor imagery electroencephalogram analysis using adaptive neural-fuzzy classification. **International Journal of Fuzzy Systems**, v. 16, n. 1, p. 111-20, Mar. 2014.
- JANG, J. R. ANFIS: adaptative-network-based fuzzy inference systems. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 23, n. 3, p. 665-685, June 1993.

JANG, J. R. **Neuro-Fuzzy and soft computing**: a computational approach to learning and machine intelligence. New York: Prentice Hall, 1997.

JASPER, H. H. The ten-twenty electrode system of the international federation. **Electroencephalography and Clinical Neurophysiology**, n. 10, p. 371-375, 1958.

KANDEL, E. R.; SCHWARTZ, J. H.; JESSELL, T. M. **Principles of neural science**. 4th ed. New York: McGraw Hill, 2000.

KROPOTOV, J. D. **Quantitative EEG, event-related potentials and neurotherapy**. San Diego: Elsevier, 2009.

LIN, C.-Y. *et al.* Combining event-related synchronization and event-related desynchronization with fuzzy C-means to classify motor imagery-induced EEG signals. INTERNATIONAL CONFERENCE ON INTELLIGENT TECHNOLOGIES AND ENGINEERING SYSTEMS, 2., 2013, Kaohsiung. **Proceedings [...]** Cham: Springer, 2014. p.1039-1045.

LUO, J.; FENG, Z.; LU, N. Spatio-temporal discrepancy feature for classification of motor imageries. **Biomedical Signal Processing and Control**, n.47, p.137-144, 2019.

MACHADO, J. C. **Pré-processamento, extração de características e classificação offline de sinais eletroencefalográficos para uso em sistemas BCI**. 2012. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.

MACHADO, J.; BALBINOT, A. Executed movement using EEG signals through a naive bayes classifier. **Micromachines**, n. 5, p. 1082-1105, 2014.

MAMDANI, E.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. **International Journal of Man-Machine Studies**, v. 7, n. 1, p. 1-13, Jan. 1975.

MANUAL EPOC. 2019. Disponível em: <https://emotiv.gitbook.io/epoc-user-manual/> Acesso em: 22 ago. 2019.

McCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115-133, 1943.

MIRNAZIRI, M. *et al.* Using Combination of μ , β and γ Bands in classification of EEG Signals. **Basic Clin Neurosci.**, v. 4, n. 1, p. 76-87, 2013.

MIRVAZIRI, H.; MOBARAKEH, Z. S. Improvement of EEG - based motor imagery classification using ring topology_based particle swarm optimization. **Biomedical Signal Processing and Control**, v. 32, p. 69-75, 2017.

MOHAMMADPOUR, M. *et al.* Comparison of EEG Signal Feature and Ensemble Learning Methods for Motor Imagery Classification. **Eighth International Conference on Information and Knowledge Technology (IKT)**, Hamedan, Iran 2016

MONTGOMERY, D. C. **Design and Analysis of Experiments**. 8th.ed. Hoboken: Wiley, 2012. 752 p. v.1, n.1.

MÜLLER, K. R. *et al.* Machine learning for real-time single-trial EEG-analysis: From brain-computer interfacing to mental state monitoring. **J. Neurosci Methodos**, v. 167, n. 1, p. 82-90, Jan. 2008.

NAKAYASHIKI, K. *et al.* Modulation of event-related desynchronization during kinematic and kinetic hand movements. **Journal of NeuroEngineering and Rehabilitation**, v. 11, p. 90-98, May 2014.

NAM, C. S. *et al.* Movement imagery-related lateralization of event-related (de)synchronization (ERD/ERS): motor-imagery duration effects. **Clinical Neurophysiology**, v. 122, n. 3, p. 567-577, Mar. 2011.

OPPENHEIM, A. V.; SCHAFER, R. W. **Discrete-time signal processing**. 2nd ed. New Jersey: Prentice-Hall, 1999.

PFURTSCHELLER, G.; SILVA, L. Event-related EEG/MEG synchronization and desynchronization: basic principles. **Clinical Neurophysiology**, v. 110, n. 11, p. 1842-1857, Nov. 1999.

RONG, H.-J. *et al.* Sequential Adaptive Fuzzy Inference System (SAFIS) for nonlinear system identification and prediction. **Fuzzy Sets and Systems**, v.157, n. 9, p.1260-1275, May 2006.

RUSSEL, S. **Artificial intelligence: a modern approach**. New Jersey: Pearson Education Limited, 2016.

SAHA, A. *et al.* Decoding of motor imagery potentials in driving using DE-induced Fuzzy-Neural Classifier. **IEEE 2nd International Conference on Recent Trends in Information Systems (ReTIS)**, p.416-421, 2015.

SANEI, S.; CHAMBERS, J. A. **EEG signal processing**. Chichester: John Wiley & Sons, 2007.

SILVA, M. G. N. M. **Redes neurais artificiais para o reconhecimento de padrões de EEG durante o movimento e imaginação do movimento das mãos**. 2014. 148 p. Tese (Doutorado em Engenharia Elétrica) - Programa de Pós-Graduação em Engenharia Elétrica, Universidade Federal de Pernambuco, Recife, 2014.

SILVERTHORN, D. U. **Fisiologia humana: uma abordagem integrada**. 5. ed. Porto Alegre: Artmed, 2010.

SOBOTTA, J. **Atlas de anatomia humana: volume tronco, vísceras e extremidade inferior**. 21. ed. Rio de Janeiro: Guanabara Koogan, 2000.

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE Transactions on Systems, man, and Cybernetics**, v. 15, n. 1, p. 116-132, Jan. 1985.

TORTORA, J. G. **Corpo humano: fundamentos da anatomia e fisiologia**. 4. ed. Porto Alegre: Artmed, 2000.

WEI, W. *et al.* Improving classification accuracy using Fuzzy Method For BCI Signal. **Bio-Medical Materials and Engineering**, n. 24, p. 2937-2943, 2014.

WELCH, P. D. The use of fast Fourier transforms for the estimation of power spectra: a method based on time averaging over short modified periodograms. **IEEE Transactions on Audio and Electroacoustics**, v. 15, n. 2, p. 70-73, June 1967.

WILLI, H. S. *et al.* **Electric machinery**. New Jersey: Word Scientific, 2015.

ZADEH, L. A. Fuzzy sets. **Information Control**, v. 8, n. 3, p. 338-353, June 1965.

APÊNDICE A

Redes Neurais Multicamadas

Redes Neurais Artificiais foram desenvolvidas inspiradas no funcionamento do Cérebro Humano, especialmente na capacidade de aprendizado, reconhecimento de padrões, solução de problemas complexos lineares, não lineares e generalização. O Cérebro Humano tem a capacidade de organizar suas estruturas neuronais, controlar seu sistema motor e realizar cálculos complexos (HAYKIN, 2009). Os algoritmos de aprendizagem de máquina são capazes de aprender um padrão associado a uma saída pré-definida e, posteriormente, prever uma saída em função de uma determinada entrada, levando em consideração este aprendizado. Os critérios de parada do aprendizado/treinamento podem estar associados com o erro quadrático médio das respostas da Rede Neural Artificial (RNA) para um determinado conjunto de dados.

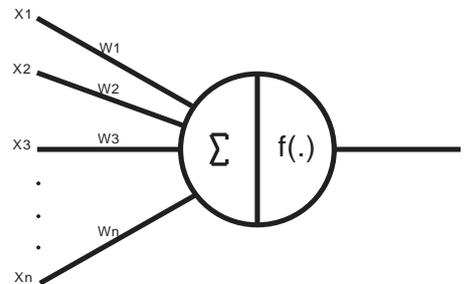
Conforme apresentado por Müller *et al.* (2008), as técnicas de aprendizagem de máquina são uma excelente escolha para compensar o comportamento estocástico dos sinais de EEG quando analisados em tempo real, pois estes sinais possuem uma variabilidade significativa de sujeito para sujeito, de ensaio para ensaio e com uma relação sinal ruído muito desfavorável, tornando-se a principal ferramenta de análise *on-line*.

Existem quatro formas de aprendizado de máquina: supervisionado, não supervisionado, semisupervisionado e aprendizagem por reforço. Aprendizagem supervisionada é quando a rede neural observa exemplos de pares de entrada e saída e aprende uma função genérica que faz o mapeamento da entrada para a saída (MÜLLER *et al.*, 2008). Na aprendizagem não supervisionada, a rede neural aprende padrões na entrada embora não tenha nenhum *feedback* explícito como, por exemplo, executando um agrupamento de dados para prever uma saída, como a rede neural apresentada por Lin *et al.* (2014). Na aprendizagem semisupervisionada, têm-se apenas poucos dados, exemplos rotulados para prever uma saída e faz-se o possível para prever as saídas não rotuladas. A aprendizagem por reforço, a rede neural aprende a partir de uma série de reforços do tipo recompensa ou punição (RUSSEL, 2016). Neste trabalho foram utilizadas apenas redes neurais com treinamento do tipo supervisionado.

O neurônio artificial (Figura 42) proposto por McCulloch e Pitts (1943), representa uma simplificação do neurônio biológico, onde as entradas (X_1 , X_2 , X_3 e X_n) representam os

dendritos, a saída (y) representa o axônio e, por fim, o corpo celular ($\Sigma|f(.)$) representa uma junção aditiva no corpo celular, é onde ocorre a função ativação do perceptron.

Figura 42 – Neurônio artificial proposto por McCulloch e Pitts.



Fonte: Adaptado livremente de McCulloch e Pitts (1943).

O neurônio idealizado por McCulloch e Pitts (1943), resolve problemas booleanos básicos linearmente separáveis, ou seja, com grupos bem definidos facilmente separáveis por uma função linear. No entanto, para problemas de uma complexidade maior, apenas um neurônio não é suficiente; uma disposição em conjuntos de neurônios faz-se necessária. Este conjunto de neurônios interligados e organizados em camadas denomina-se Redes Neurais Multicamadas (RUSSEL, 2016). As interconexões entre os neurônios da rede representam as sinapses (pontos de contato entre os neurônios cerebrais) onde as saídas (axônios) de cada neurônio encontram-se conectadas na entrada (dendrito) do próximo neurônio. Essa interconexão possui um peso associado representado pelo parâmetro w_i ($w_1, w_2, w_3 \dots w_n$), as sinapses podem ser excitatórias ou inibitórias, portanto, devido à interconexão entre a entrada e a saída dos neurônios de cada camada da rede neural, a entrada é multiplicada pelo peso resultando em $x_i w_i$. No perceptron ocorre uma soma ponderada entre os valores das n entradas e uma comparação com a função de ativação do perceptron, como apresentado na Equação (19).

$$net = \sum_{i=0}^n x_i w_i \quad (19)$$

A saída y é dada pela função ativação que é definida a partir dos vetores de peso e de entrada, ou seja, o valor da soma ponderada é o argumento da função de ativação. Existem diversos tipos básicos de função ativação mais utilizados que podem ser consultados em Haykin (2009).

O valor de saída do neurônio considerando todas as variáveis (entradas, pesos e função de ativação), pode ser calculado pela Equação (20):

$$Saída = y(net) \quad (20)$$

Para problemas não linearmente separáveis, utilizam-se redes neurais com uma ou mais camadas intermediárias de neurônios. A função destes neurônios das camadas intermediárias é extrair informações dos neurônios das camadas de entrada, processar a informação e fornecê-las como entrada para os neurônios da camada de saída (BRAGA *et al.*, 2007). Este tipo de interconexão é chamado de rede *feedforward* (propagação para frente), pois cada camada se conecta com a próxima camada sem ocorrer uma realimentação entre as entradas e saídas das camadas (GARDNER; DORLING, 1998).

Para possibilitar a aprendizagem, ou seja, a correção do erro, de uma rede *feedforward* o algoritmo *backpropagation* pode ser utilizado conforme apresentado por Gardner e Dorling (1998). Esta técnica de treinamento realiza uma propagação do erro no sentido contrário ao do sinal, ou seja, da saída em direção à entrada, atualizando os pesos com o objetivo de reduzir o erro até que o erro mínimo seja obtido.

O algoritmo *backpropagation* é um método baseado no gradiente descendente com uma complexidade computacional ocasionada pelo cálculo da derivada parcial, porém este método converge linearmente com uma velocidade muito lenta. No entanto, o algoritmo Levenberg-Marquardt (LM) apresenta uma resposta mais rápida, utilizando o método gradiente descendente combinado com a convergência do método de Gauss-Newton (CHEN; ZHANG, 2012).

Considerando a otimização de uma função de segunda ordem $F(w)$ e seja g seu vetor de gradiente e H seu vetor Hessiano. De acordo com o método de Levenberg-Marquardt, o ajuste ideal Δw aplicado ao vetor de parâmetro w pode ser definido pela Equação (21):

$$\Delta w = [H + \lambda I]^{-1} g \quad (21)$$

onde I é a matriz identidade com as mesmas dimensões que H e λ é o parâmetro regularizador que força a matriz soma $[H + \lambda I]$ a ser positiva. Com base neste conhecimento e considerando um rede multicamada com um Perceptron de saída. A rede é treinada minimizando a função de custo pela Equação (22):

$$E_{av}(w) = \frac{1}{2N} \sum_{i=1}^N [d(i) - F(x(i); w)]^2 \quad (22)$$

onde $\{X(i), d(i)\}_{i=1}^N$ representa a amostra de treinamento e $F(x(i); w)$ é a função aproximação realizada pela rede; os pesos sinápticos da rede são organizados de maneira ordenada para formar o vetor de peso w . O Gradiente Hessiano da função custo E_{av} é definido respectivamente pelas Equações (23) e (24):

$$g(w) = \frac{\partial E_{av}(w)}{\partial w} \quad (23)$$

$$H(w) = \frac{\partial^2 E_{av}(w)}{\partial w^2} \quad (24)$$

Então substituindo a Equação (23) e a Equação (24) na Equação (21), então o ajuste do parâmetro Δw é calculado para cada iteração do algoritmo de Levenberg-Marquardt. No entanto quando a dimensionalidade do vetor de peso w for alta a complexidade da Equação (21) aumenta. A complexidade computacional é atribuída a natureza complexa da matriz Hessiana ($H(w)$). Para atenuar essa dificuldade, o procedimento recomendado é ignorar o segundo termo no lado direito da Equação (21) aproximando a matriz Hessiana pela Equação (25):

$$H(w) \approx \frac{1}{N} \sum_{i=1}^N \left[\frac{\partial F(X(i); w)}{\partial w} \right] \left[\frac{\partial F(x(i); w)}{\partial w} \right]^T \quad (25)$$

Essa aproximação é reconhecida como produto externo da derivada parcial $\partial F(w, x(i))/\partial w$ calculado pela média da amostra de treinamento; portanto, é referida como aproximação do produto externo da Hessians. O uso desta aproximação é justificado quando o algoritmo Levenberg-Marquardt está operando na vizinhança de um mínimo local ou global. A aproximação do algoritmo de Levenberg-Marquardt é um método adequado para problemas de estimativa de mínimo quadrados não lineares.

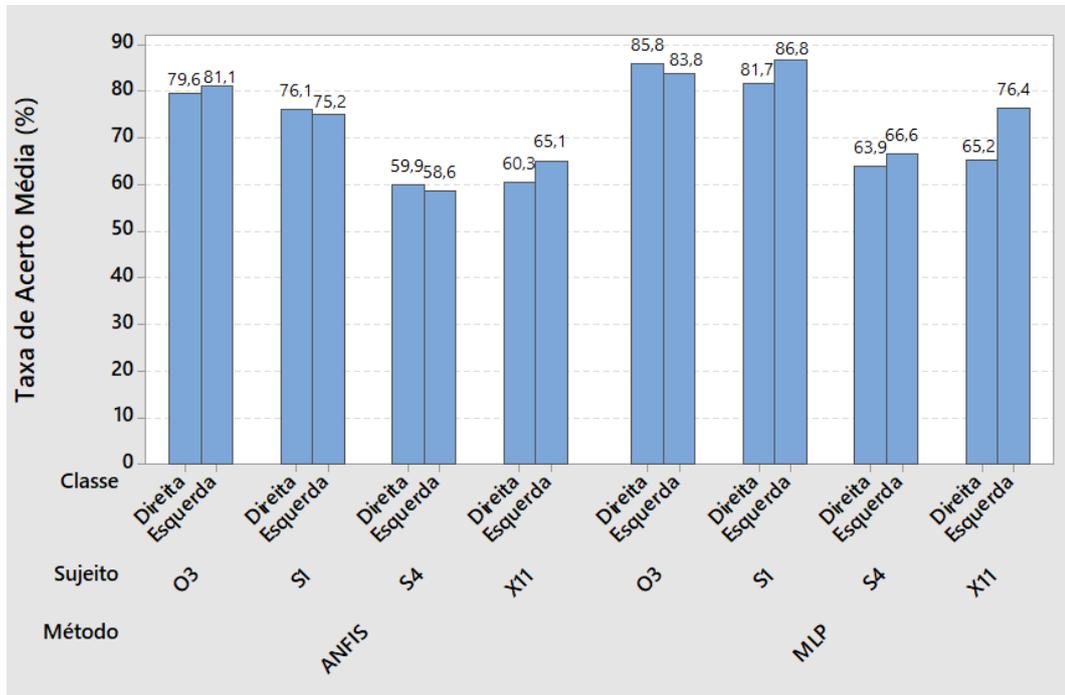
O parâmetro regularizador λ desempenha um papel crítico no funcionamento do algoritmo Levenberg-Marquardt. Se reduzirmos este parâmetro para zero, a Equação (21) reduz-se ao método de Newton. Em contrapartida se atribuirmos um valor muito alto para o parâmetro λ de tal forma que o λI seja predominante em relação à Hessiana, o algoritmo Levenberg-Marquardt funcionará efetivamente com o método do gradiente descendente. A cada iteração do algoritmo, o valor atribuído a este parâmetro deve ser grande o suficiente para manter a matriz soma $(H + \lambda I)$ positiva. (HAYKIN, 2009; PRESS *et al.*, 1988).

Os passos para desenvolver o algoritmo LM são:

1. calcular o E_{av} na iteração $n - 1$
2. escolher um valor pequeno para o parâmetro λ , sugere-se: $\lambda = 10^{-3}$;
3. solucionar a Equação (21), ajustando Δw na iteração n e avaliar $E_{av}(w + \Delta w)$;
4. se $E_{av}(w + \Delta w) \geq E_{av}$, incrementa λ e retorna para o passo 3;
5. se, por outro lado, $E_{av}(w + \Delta w) < E_{av}$, decrementa λ e atualiza a solução teste $w \rightarrow w + \Delta w$, então retorna para o passo 3. ;

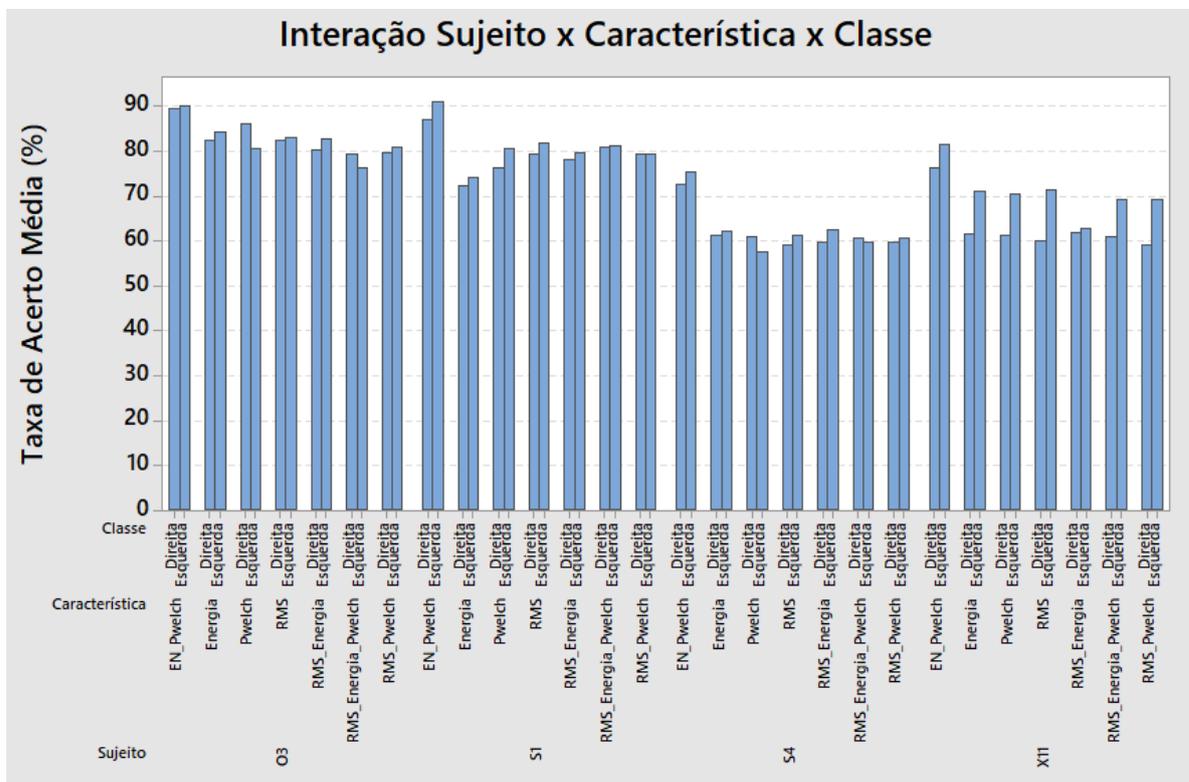
ANEXO A

Figura 43 – Gráfico da Interação Método x Sujeito x Classe.



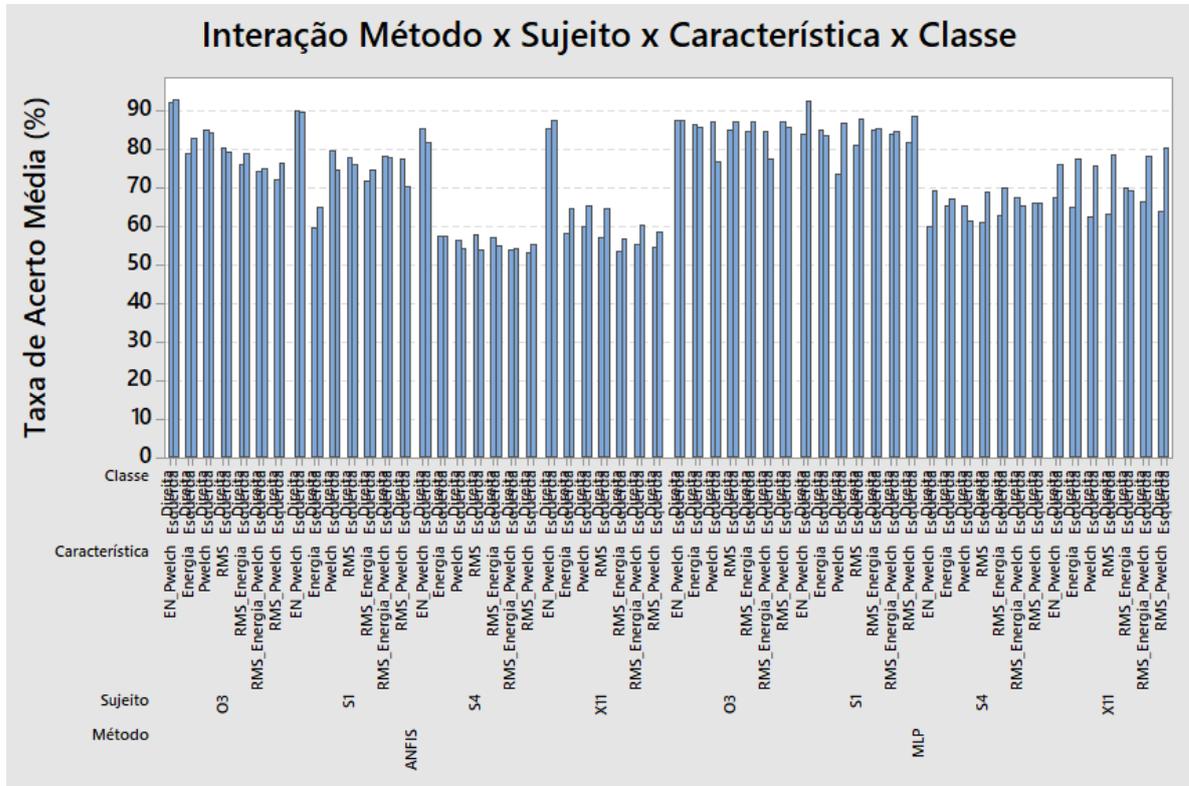
Fonte: Autor

Figura 44 – Gráfico da Interação Sujeito x Característica x Classe.



Fonte: Autor

Figura 45 – Gráfico da Interação Método x Sujeito x Característica x Classe.



Fonte: Autor