

**UNIVERSIDADE DE SÃO PAULO**  
Instituto de Ciências Matemáticas e de Computação

## Seleção de canais para BCIs baseadas no P300

**Pedro Henrique da Costa Ulisses**

Dissertação de Mestrado do Programa de Pós-Graduação em Ciências de Computação e Matemática Computacional (PPG-C<sup>2</sup>MC)



SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: \_\_\_\_\_

**Pedro Henrique da Costa Ulisses**

## Seleção de canais para BCIs baseadas no P300

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestre em Ciências – Ciências de Computação e Matemática Computacional. *VERSÃO REVISADA*

Área de Concentração: Ciências de Computação e Matemática Computacional

Orientador: Prof. Dr. João Luís Garcia Rosa

**USP – São Carlos**  
**Março de 2019**

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi  
e Seção Técnica de Informática, ICMC/USP,  
com os dados inseridos pelo(a) autor(a)

U39s           Ulisses, Pedro Henrique da Costa  
                Seleção de canais para BCIs baseadas no P300 /  
                Pedro Henrique da Costa Ulisses; orientador João  
                Luís Garcia Rosa. -- São Carlos, 2018.  
                81 p.

                Dissertação (Mestrado - Programa de Pós-Graduação  
                em Ciências de Computação e Matemática  
                Computacional) -- Instituto de Ciências Matemáticas  
                e de Computação, Universidade de São Paulo, 2018.

                1. BCI. 2. Ensemble de classificadores. 3. P300.  
                4. Seleção de canais. 5. Máquina de Vetor de  
                Suporte. I. Luís Garcia Rosa, João , orient. II.  
                Título.

**Pedro Henrique da Costa Ulisses**

## Channel selection for P300-based BCIs

Master dissertation submitted to the Institute of Mathematics and Computer Sciences – ICMC-USP, in partial fulfillment of the requirements for the degree of the Master Program in Computer Science and Computational Mathematics. *FINAL VERSION*

Concentration Area: Computer Science and Computational Mathematics

Advisor: Prof. Dr. João Luís Garcia Rosa

**USP – São Carlos**  
**March 2019**



*À minha mãe Vera Lúcia e  
ao meu avô Nestor Barbosa (in memoriam).*





# AGRADECIMENTOS

---

---

Ao professor Dr. João Luís Garcia Rosa, pela orientação e paciência ao longo desse projeto. Ao CNPq pelo apoio financeiro para a realização desta pesquisa. Agradecimentos especiais aos meus irmãos George Augusto e Verônica Ulisses, às minhas tias Bete, Carmem, Coca e Graça, à minha namorada Anne Kareninne e às minhas amigas Amanda Rêgo e Denise Dália.



*“É a história dum Davi contra um Golias. O pequeno Davi da Bíblia venceu o gigante Golias apenas com uma pedra e uma funda. Mas há outros meios de se derrubar grandes obstáculos. A persistência não é o mais prático mas talvez seja de todos o mais eficiente.”*  
*(Marcos Rey)*



# RESUMO

ULISSES, P. H. C. **Seleção de canais para BCIs baseadas no P300**. 2019. 81 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Interface Cérebro-Computador é um meio que permite a comunicação do cérebro com dispositivos externos e tem como principal público-alvo as pessoas com problemas motores, incapazes de se comunicarem e/ou se locomoverem. Uma das principais aplicações são os soletradores baseados no P300 que fornecem um meio de indivíduos se comunicarem através de um teclado virtual. Devolver a capacidade de comunicação para uma pessoa é de extrema importância para a qualidade de vida das pessoas. Esse tipo de aplicação possui diversos desafios, um deles é a necessidade da BCI ser treinada especificamente para cada indivíduo. Esse treinamento pode levar horas e até mesmo dias. Uma das formas de diminuir esse tempo é utilizar um dos conjuntos de canais pré-definidos que são sugeridos na literatura, porém esses conjuntos não garantem um funcionamento adequado da BCI, o que pode frustrar os indivíduos não desejarem mais utilizar uma BCI. Para solucionar esse problema, é proposto no presente trabalho a seleção de canais a partir de um conjunto de canais para agilizar o processo de treinamento e atingir um ótimo desempenho com a BCI.

**Palavras-chave:** BCI, *Ensemble* de Classificadores, P300, Seleção de Canais, Máquina de Vetor de Suporte.



# ABSTRACT

ULISSES, P. H. C. **Channel selection for P300-based BCIs**. 2019. 81 p. Dissertação (Mestrado em Ciências – Ciências de Computação e Matemática Computacional) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2019.

Brain-Computer Interface is a means that allows the communication of the brain with external devices and has as main target audience the people with motor problems, unable to communicate and/or move around. One of the main applications is the P300-based spellers that provide a means for individuals to communicate through a virtual keyboard. Recovering the ability to communicate to a person is of extreme importance to the quality of people's lives. This type of application has several challenges, one of which is the need for BCI to be trained specifically for each individual. This training can take hours and even days. One of the ways to decrease this time is to use one of the predefined set of channels that are suggested in the literature, but these sets do not guarantee an adequate functioning of BCI, which can frustrate individuals no longer want to use a BCI. To solve this problem, it is proposed in the present work the selection of channels from a set of channels to accelerate the training process and achieve optimal performance with BCI.

**Keywords:** BCI, *Ensemble* of Classifiers, P300, Channel Selection, Support Vector Machine.





# LISTA DE ILUSTRAÇÕES

---

---

Figura 1 – Sistema de BCI básico. Adaptado de (WOLPAW <i>et al.</i> , 2002). . . . .	31
Figura 2 – Posição dos eletrodos no sistema 10-20. Uso autorizado pelo autor (RAMA-DAN <i>et al.</i> , 2015). . . . .	31
Figura 3 – Demonstração do paradigma <i>oddball</i> . . . . .	35
Figura 4 – Média dos sinais com e sem o P300. . . . .	35
Figura 5 – Hiperplano da SVM para separar as classes maximizando a margem $\gamma$ . Adaptado de (KAPER <i>et al.</i> , 2004). . . . .	44



# LISTA DE ALGORITMOS

---

---

Algoritmo 1 – Algoritmo de Seleção de Canais . . . . .	42
--	----



# LISTA DE TABELAS

---

---

Tabela 1 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de SVMs com 64 canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,1 dB. . . . .	53
Tabela 2 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de SVMs com 64 canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,5 dB. . . . .	54
Tabela 3 – Avaliação de cada modelo para o indivíduo A com o aumento do número de canais através da seleção de canais com ondulação máxima 0,1 dB. . . . .	55
Tabela 4 – Avaliação de cada modelo para o indivíduo B com o aumento do número de canais através da seleção de canais com ondulação máxima 0,1 dB. . . . .	55
Tabela 5 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de SVMs após seleção de canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,1 dB. . . . .	56
Tabela 6 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de SVMs após seleção de canais e aumentando o número de sequências e utilizando dados dos indivíduos A e B com ondulação máxima de 0,5 dB. . . . .	56
Tabela 7 – Avaliação dos classificadores para o indivíduo A com o aumento do número de canais através da seleção de canais com ondulação máxima de 0,5 dB. . . . .	57
Tabela 8 – Avaliação dos classificadores para o indivíduo B com o aumento do número de canais através da seleção de canais com ondulação máxima de 0,5 dB. . . . .	58
Tabela 9 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de 17 SVMs e aumentando o número de sequências. . . . .	58
Tabela 10 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de 17 SVMs após seleção de canais, iniciando com os canais Pz, Oz, PO7 e PO8 fixos, e aumentando o número de sequências utilizando dados dos indivíduos A e B. . . . .	59
Tabela 11 – Avaliação dos <i>ensembles</i> com 17 SVMs para o indivíduo A com a seleção de canais iniciando com os canais Pz, Oz, PO7 e PO8 selecionados. . . . .	59
Tabela 12 – Avaliação dos <i>ensembles</i> com 17 SVMs para o indivíduo B com a seleção de canais iniciando com os canais Pz, Oz, PO7 e PO8 selecionados. . . . .	60
Tabela 13 – Número de caracteres classificados corretamente por um <i>ensemble</i> de 17 SVMs (RAKOTOMAMONJY; GUIGUE, 2008). . . . .	61
Tabela 14 – Quantidade de vezes que cada canal foi selecionado utilizando <i>ensemble</i> com 17 SVMs para o indivíduo A. . . . .	62

Tabela 15 – Quantidade de vezes que cada canal foi selecionado utilizando <i>ensemble</i> com 17 SVMs para o indivíduo B. . . . .	62
Tabela 16 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de 17 LDAs utilizando diferentes combinações de canais. . . . .	63
Tabela 17 – Quantidade de vezes que cada canal foi selecionado utilizando 17 LDA's para o indivíduo A. . . . .	63
Tabela 18 – Quantidade de vezes que cada canal foi selecionado utilizando 17 LDA's para o indivíduo B. . . . .	64
Tabela 19 – Número médio de canais selecionados por cada um dos 17 modelos dos <i>ensembles</i> de SVMs e LDAs. . . . .	64
Tabela 20 – Resultados obtidos por <i>ensembles</i> com modelos ponderados. . . . .	65
Tabela 21 – Resultados obtidos por <i>ensembles</i> de 17 SVMs. . . . .	65
Tabela 22 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de 5 SVMs com diferentes combinações de canais e aumentando o número de seqüências. . . . .	66
Tabela 23 – Quantidade de vezes que cada canal foi selecionado utilizando 5 SVMs para o indivíduo A. . . . .	67
Tabela 24 – Quantidade de vezes que cada canal foi selecionado utilizando 5 SVMs para o indivíduo B. . . . .	67
Tabela 25 – Número de caracteres classificados corretamente pelo <i>ensemble</i> de 5 LDAs utilizando diferentes combinações de canais. . . . .	68
Tabela 26 – Quantidade de vezes que cada canal foi selecionado com 5 LDAs para o indivíduo A. . . . .	69
Tabela 27 – Quantidade de vezes que cada canal foi selecionado com 5 LDAs para o indivíduo B. . . . .	70
Tabela 28 – Número médio de canais selecionados por cada um dos 5 modelos dos <i>ensembles</i> de SVMs e LDAs. . . . .	70
Tabela 29 – Número de caracteres classificados corretamente por uma única SVM utilizando diferentes combinações de canais. . . . .	71
Tabela 30 – Quantidade de vezes que cada canal foi selecionado utilizando uma única SVM para o indivíduo B. . . . .	72
Tabela 31 – Número de caracteres classificados corretamente por uma única LDA utilizando diferentes combinações de canais. . . . .	72
Tabela 32 – Quantidade de vezes que cada canal foi selecionado utilizando uma única LDA para o indivíduo A. . . . .	73
Tabela 33 – Quantidade de vezes que cada canal foi selecionado utilizando uma única LDA para o indivíduo B. . . . .	73
Tabela 34 – Número de canais selecionados pela SVM e LDA. . . . .	73

---

# LISTA DE ABREVIATURAS E SIGLAS

---

---

BCI	<i>Brain-Computer Interface</i>
BSS	<i>Blind Source Separation</i>
ECG	Eletrocardiograma
EEG	Eletroencefalograma
ELA	Esclerose Lateral Amiotrófica
EMG	Eletromiograma
EOG	Eletrooculograma
ERD	<i>Event-Related Desynchronization</i>
ERP	<i>Event Related Potential</i>
ERS	<i>Event-Related Synchronization</i>
EWLDA	<i>Ensemble Weighted of Linear Discriminant Analysis</i>
EWSVM	<i>Ensemble Weighted of Support Vector Machine</i>
fMRI	<i>Functional Magnetic Resonance Imaging</i>
LDA	<i>Linear Discriminant Analysis</i>
LIS	<i>Locked-in Syndrome</i>
MEG	Magnetoencefalografia
MI	<i>Motor Imagery</i>
PCA	<i>Principal Component Analysis</i>
PET	<i>Positron Emission Tomography</i>
PSO	<i>Particle Swarm Optimization</i>
REM	<i>Rapid Eye Movement</i>
RNA	Redes Neurais Artificiais
SCP	<i>Slow Cortical Potentials</i>
SNR	<i>Signal to Noise Ratio</i>
SVM	<i>Support Vector Machine</i>
VEP	<i>Visual-Evoked Potentials</i>





# SUMÁRIO

---

---

1	<b>INTRODUÇÃO</b>	25
1.1	Objetivos e Hipótese	26
1.2	Organização do Trabalho	26
2	<b>INTERFACE CÉREBRO-COMPUTADOR</b>	29
2.1	Considerações Iniciais	29
2.2	Brain-Computer Interface	29
2.2.1	<i>Componentes da BCI</i>	30
2.2.2	<i>Eletroencefalograma</i>	32
2.3	<i>Event-Related Potentials</i>	33
2.3.1	<i>Componentes do ERP</i>	33
2.3.2	<i>Componente P300</i>	34
2.4	Considerações Finais	35
3	<b>EXTRAÇÃO, SELEÇÃO E CLASSIFICAÇÃO</b>	37
3.1	Considerações Iniciais	37
3.2	Remoção de Artefatos e Ruídos	37
3.3	Extração de Características	38
3.3.1	<i>Filtro Chebyshev Tipo I</i>	38
3.3.2	<i>Normalização</i>	39
3.3.3	<i>Redução de Dimensionalidade</i>	39
3.4	Seleção de Canais	40
3.5	Classificadores	42
3.6	Considerações Finais	47
4	<b>METODOLOGIA E RESULTADOS</b>	49
4.1	Considerações Iniciais	49
4.2	Descrição da Base de dados	49
4.3	Pré-processamento dos Dados	50
4.4	Seleção de Características	51
4.4.1	<i>Filtragem</i>	51
4.4.2	<i>Redução da dimensionalidade</i>	51
4.5	<i>Ensemble de Classificadores</i>	51
4.6	Experimentos e Resultados	52

<b>4.6.1</b>	<b>17 Classificadores</b> . . . . .	<b>53</b>
<b>4.6.2</b>	<b>5 Classificadores</b> . . . . .	<b>65</b>
<b>4.6.3</b>	<b>1 Classificador</b> . . . . .	<b>70</b>
<b>4.7</b>	<b>Considerações Finais</b> . . . . .	<b>74</b>
<b>5</b>	<b>CONCLUSÕES</b> . . . . .	<b>75</b>
<b>5.1</b>	<b>Contribuições e Limitações</b> . . . . .	<b>75</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>77</b>

---

# INTRODUÇÃO

---

Interface cérebro-computador (*Brain-Computer Interface* - BCI) é um sistema que permite a comunicação do cérebro com dispositivos externos sem o uso de nervos periféricos e de músculos (MILLÁN *et al.*, 2010). Essa comunicação é realizada com a aquisição dos sinais da atividade cerebral de um indivíduo, processamento dos sinais e tradução dos sinais processados em um ou mais comandos que são enviados para um dispositivo ou aplicação. A aquisição do sinal na maioria das vezes é de forma não invasiva por ser considerada a forma mais segura. Há vários métodos não-invasivos existentes, tais como Imagem por Ressonância Magnética Funcional (fMRI), Tomografia por Emissão de Pósitrons (PET), Magnetoencefalografia (MEG) e a Eletroencefalografia que é a mais utilizada nas aplicações de BCI por causa da sua alta resolução temporal, portabilidade e baixo custo (MARSHALL *et al.*, 2013). O registro da atividade cerebral é chamado de Eletroencefalograma (EEG). Na etapa de processamento do sinal há remoção do ruído e realizada a extração de características do sinal, em seguida os sinais são utilizados para treinar um modelo computacional para traduzir novos sinais de entrada em comandos que serão utilizados em uma aplicação (RAMADAN *et al.*, 2015).

As BCIs podem ser utilizadas em diversos tipos de aplicações, tanto para indivíduos com e sem problemas motores ou neurológicos. Dentre as aplicações, podemos destacar os soletradores, controles para próteses, controles para aparelhos domésticos, controle de cadeira de rodas, tratamento do déficit de atenção e jogos virtuais. Mesmo com o avanço das pesquisas em BCI e ela sendo utilizada nas mais diversas áreas, o foco principal desta área é o desenvolvimento de aplicações voltadas para os indivíduos que possuem problemas neurológicos ou musculares severos, tais como os indivíduos que sofreram acidente vascular no tronco cerebral, paralisia cerebral, lesão na coluna espinhal (LEZA; PUTHUSSERYPADY, 2017) e indivíduos que sofrem com a síndrome do encarceramento (*Locked-In Syndrome* - LIS). A LIS caracteriza-se basicamente por tetraplegia, anartria e preservação do nível de consciência, além de certa movimentação ocular pela qual o indivíduo se comunica (NIJBOER *et al.*, 2008).

As aplicações de soletradores que utilizam uma BCI para controlá-las, são as típicas aplicações desenvolvidas para fornecer uma alternativa aos indivíduos, que sofrem de LIS, de se comunicarem. Por esses indivíduos possuírem a capacidade de controlar o movimento dos olhos, as BCIs baseadas no P300 são uma opção bastante utilizada (NIJBOER *et al.*, 2008). O P300 é um componente presente na atividade cerebral que ocorre após o indivíduo ser apresentado a um estímulo raro em uma sequência de estímulos frequentes. Porém, assim como todas as aplicações de BCIs, as baseadas no P300 também possuem diversos desafios, tais como: os sinais coletados têm pequena amplitude, necessitando de amplificadores, há bastante ruído, isso leva a uma alta taxa de erro que provoca a classificação de sinais imprecisa, a taxa de transferência de dados é baixa, prejudicando o desempenho das aplicações (RAMADAN *et al.*, 2015).

As aplicações de BCI geralmente precisam ser calibradas para cada indivíduo, devido haver diferença na atividade cerebral entre os indivíduos. E esta etapa visa melhorar o desempenho da BCI buscando as melhores configurações, tais como um conjunto de canais que melhorem o desempenho da aplicação (SPEIER; DESHPANDE; POURATIAN, 2015). A diminuição do número de canais é interessante pois propicia a redução da quantidade de dados que serão utilizados. Isso vai permitir diminuir o tempo de treinamento do classificador, evitar o problema da dimensionalidade, permitir que os capacetes/toucas para gravação do EEG possuam poucos canais e assim agilizar o posicionamento deles no indivíduo. Há trabalhos que indicam os melhores canais que podem ser utilizados nas BCIs, porém estes canais nem sempre garantem um bom desempenho para os usuários, por isso a importância da seleção de canais.

## 1.1 Objetivos e Hipótese

Este projeto tem como hipótese que a seleção de canais proposta irá reduzir o número de canais sem prejudicar o desempenho. Como objetivo geral, diminuir o tempo de treinamento de um classificador de uma BCI baseada no P300 através da seleção de canais. E como objetivos específicos:

- Refinar os parâmetros da etapa de extração de características dos sinais de EEG e dos modelos computacionais;
- Variar os classificadores utilizados, *ensemble* com Máquina de Vetores de Suporte e Análise Discriminante Linear, e comparar o desempenho de cada um deles com os resultados obtidos em outros trabalhos que utilizaram a mesma base de dados.

## 1.2 Organização do Trabalho

Além deste capítulo introdutório, este texto de dissertação de mestrado está estruturado em mais três capítulos, que estão organizados da seguinte forma:

- Capítulo 2: Interface Cérebro-Computador, onde são apresentados alguns conceitos de BCI.
- Capítulo 3: Extração, Seleção e Classificação de Características, apresenta as etapas de uma BCI e algumas técnicas que podem ser utilizadas fazendo o paralelo com as adotadas no presente trabalho.
- Capítulo 4: Metodologia e Resultados, detalha todas as etapas realizadas nos testes desenvolvidos neste trabalho, desde o processamento dos sinais até a avaliação do desempenho dos modelos desenvolvidos e discute os resultados obtidos, comparando com resultados da literatura.
- Capítulo 5: Conclusão, onde é apresentado de forma resumida o que foi feito no trabalho, os objetivos atingidos, contribuições, limitações e continuação do trabalho.



---

# INTERFACE CÉREBRO-COMPUTADOR

---

## 2.1 Considerações Iniciais

Nesse capítulo são apresentados os conceitos fundamentais para o entendimento do presente projeto de mestrado. Dessa forma, será apresentada uma breve história sobre BCI e soletradores, descrição do funcionamento de uma BCI, descrição do P300 e das BCIs baseadas no P300.

## 2.2 Brain-Computer Interface

Interface cérebro-computador (*Brain-computer interface* - BCI) é um sistema que permite um indivíduo controlar um dispositivo através de sinais da sua atividade cerebral que são adquiridos de várias regiões do cérebro. Em 1970, pesquisadores desenvolveram um sistema primitivo de controle baseado na atividade cerebral adquirida do couro cabeludo ([VALLABHANENI; WANG; HE, 2005](#)); foi a primeira vez que o termo BCI foi citado na literatura científica, porém a ideia de conectar o cérebro a outros dispositivos não era novidade ([VIDAL, 1973](#)). Testes com sinais do córtex motor de macacos já haviam sido realizados antes, em um trabalho de Fetz publicado em fevereiro de 1969, e pesquisadores demonstraram que os comportamentos naturais de aprendizagem do cérebro poderiam adaptar-se aos novos estímulos, bem como controlar a ativação de áreas específicas do cérebro ([FETZ, 1969](#)).

No ano de 1977, Jacques Vidal, pesquisador do *Brain Computer Interface Laboratory* da Universidade da Califórnia de Los Angeles (UCLA), demonstrou que as respostas a estímulos visuais em humanos podem ser utilizadas como um canal de comunicação. Os sinais de EEG foram utilizados para controlar um cursor em um labirinto 2D ([VIDAL, 1977](#)).

Os pesquisadores [Farwell e Donchin \(1988\)](#) desenvolveram um soletrador que mostrava uma matriz, com 6 colunas e 6 linhas, contendo 26 letras e opções adicionais como inserir espaço,

apagar caractere e sair. A aplicação funcionava baseada no paradigma *oddball*, que é uma forma de induzir mudanças na atividade cerebral apresentando dois tipos de evento, sendo um deles frequente e o outro raro. Quando o evento raro acontece, provoca alterações na atividade cerebral que é reconhecida pela BCI, e dessa forma é possível identificar o comando focado pelo usuário. Nesta aplicação o evento raro era a intensificação de uma linha ou coluna, o evento frequente era o estado inicial da linha ou da coluna. A alteração na atividade cerebral ocorre quando a linha ou a coluna que contém a letra que o usuário está focando é intensificada. O componente do sinal que é analisado em aplicações baseadas no paradigma *oddball* é o P300, que será discutido neste capítulo.

Mais tarde, um grupo de pesquisa liderado por Gert Pfurtscheller desenvolveu um teclado virtual controlado por uma BCI baseada em imagética motora para pessoas que sofrem de esclerose lateral amiotrófica (ELA). Imagética motora é a imaginação de uma ação motora sem que ocorra o movimento de fato (PFURTSCHELLER; NEUPER, 2001). Os pacientes que testaram essa aplicação passaram por mais de 100 sessões de treinamento que consistiam em olhar para um monitor e imaginar o movimento da mão direita ou esquerda de acordo com uma seta que aparecia no monitor para um dos lados. O *feedback* para o usuário era baseado em uma barra no monitor que aumentava de tamanho para o lado identificado pela BCI. Para o usuário escrever uma palavra, ele tem que escolher uma dentre duas opções utilizando imagética motora das mãos, ou seja, imaginar mover uma das mãos. Cada opção corresponde a um conjunto de letras e cada vez que um conjunto é selecionado ele é dividido ao meio; esses passos são realizados até que o conjunto possua apenas uma letra. Após a letra ser selecionada o usuário tem a opção de utilizá-la na palavra ou descartá-la. Esta aplicação permite os usuários escreverem aproximadamente uma letra em dois minutos, sendo considerada lenta em comparação com outros dispositivos de entrada (teclado e mouse), porém fornece um meio de comunicação para pessoas com ELA e outras deficiências neuromotoras (OBERMAIER; MULLER; PFURTSCHELLER, 2003).

### 2.2.1 Componentes da BCI

A BCI possui um ciclo entre o usuário e o sistema. Geralmente após o usuário interagir com o sistema, através da atividade cerebral, um *feedback* é apresentado ao usuário sobre a interação ocorrida. O ciclo da BCI é composto pelas seguintes etapas: aquisição do sinal; pré-processamento; extração de características; tradução das características e execução do comando (WOLPAW *et al.*, 2002). Na Figura 1 é apresentado um esquema simplificado de uma BCI.

Na etapa de aquisição do sinal, o indivíduo interage com a aplicação e a atividade cerebral é gravada por meio de eletrodos que são posicionados no couro cabeludo, no caso do EEG. Posteriormente o sinal é amplificado, digitalizado e enviado para um computador para em seguida ser pré-processado (WOLPAW *et al.*, 2002; HE *et al.*, 2013). O posicionamento dos eletrodos geralmente segue o padrão do sistema 10-20 e normalmente são utilizados 32, 64, 128 ou 256 eletrodos. A Figura 2 apresenta o sistema 10-20. Atualmente existem toucas



que auxiliam no posicionamento dos eletrodos e também há *headsets* com os eletrodos já nas posições adequadas (FREEMAN; QUIAN-QUIROGA, 2012).

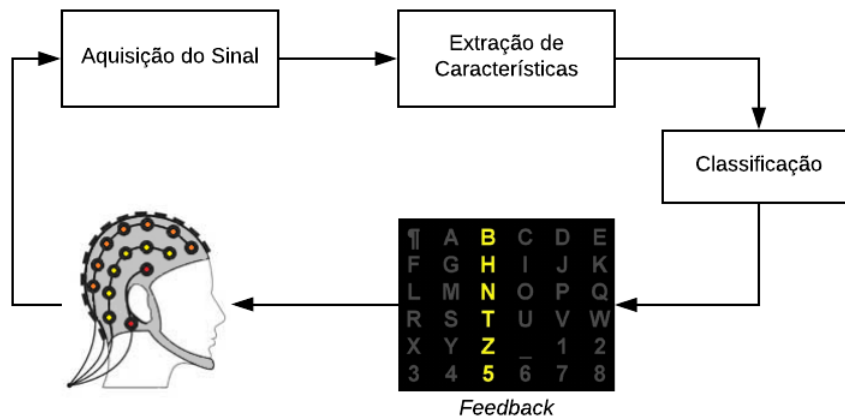


Figura 1 – Sistema de BCI básico. Adaptado de (WOLPAW *et al.*, 2002).

Durante a etapa de seleção e extração de características, o sinal é submetido a um ou mais processos, tais como: medidas da amplitude, análise espectral ou filtros espaciais. Características eficientes têm uma forte correlação com a intenção do usuário. As características comumente extraídas dos sinais de EEG estão no domínio do tempo ou da frequência. Para melhorar a acurácia das características extraídas, os artefatos podem ser removidos com o auxílio do eletromiograma (EMG) dos músculos do crânio e do eletrooculograma (EOG).

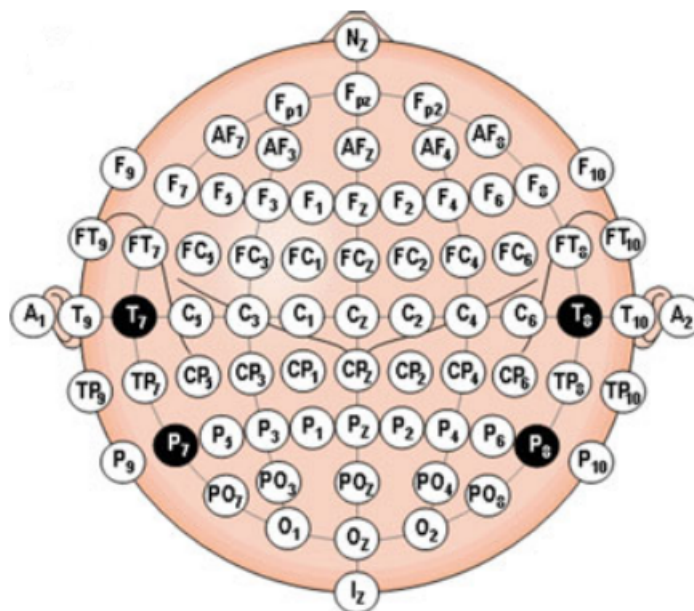


Figura 2 – Posição dos eletrodos no sistema 10-20. Uso autorizado pelo autor (RAMADAN *et al.*, 2015).

Na etapa de tradução do sinal, as características extraídas são fornecidas como entrada para um algoritmo tradutor, que é um algoritmo de classificação previamente treinado que con-

verte as entradas em um comando. O treinamento prévio pode ser realizado com características extraídas de um indivíduo durante sessões em que ele utilizou a BCI com o objetivo de ajustá-la. Outra possibilidade é do algoritmo ter sido treinado previamente com características de vários indivíduos e quando um novo indivíduo for utilizar a BCI não há a necessidade do treinamento prévio, o algoritmo tradutor é capaz de se adaptar ao indivíduo à medida que ele utiliza a BCI.

Devido a atividade cerebral sofrer alterações de curto e longo prazos devido a diversos fatores (fadiga, doença, ambiente), o algoritmo tradutor deve ser capaz de acompanhar e adaptar-se a mudanças nas características do sinal, de modo a fornecer um controle eficaz e mais eficiente possível. A etapa de treinamento, que algumas aplicações de BCI possuem, tem o objetivo de treinar este algoritmo para reconhecer os padrões das entradas durante o teste da aplicação.

A saída da BCI é o comando que o algoritmo de tradução produziu, e na maioria das vezes ocorre na tela do computador. Esta saída representa a intenção do usuário, selecionando uma letra no soletrador, movendo um objeto na tela do computador ou movendo uma cadeira de rodas. Este comando é considerado o *feedback* do sistema que o cérebro usa para melhorar o desempenho do usuário. Após a execução do comando uma nova interação do indivíduo com a BCI acontece e um novo ciclo se inicia (WOLPAW *et al.*, 2002; HE *et al.*, 2013).

### 2.2.2 Eletroencefalograma

Eletroencefalograma (EEG) é a atividade elétrica registrada a partir da superfície do couro cabeludo através de eletrodos que são sensíveis às mudanças de potenciais pós-sinápticos de neurônios no córtex cerebral. O EEG é adquirido através do processo, anteriormente citado, conhecido como eletroencefalografia. Os eletrodos possuem uma sensibilidade padrão e medem os potenciais na área próxima a ele; esses potenciais elétricos são combinados e amplificados para gerar uma atividade rítmica. Essa atividade é classificada em bandas de frequência: delta, teta, alfa, beta, mu e gama (SZAFIR, 2010; MOAZZAMI, 2012).

A banda delta está no intervalo de 0,1-3,5 Hz, tende a ser a amplitude mais alta e a onda mais lenta, está relacionada ao sono em adultos e bebês. A banda teta está no intervalo de 3,5-7,5 Hz, e está associada a movimentos oculares rápidos (*rapid eye movement* - REM), sono, resolução de problemas, atenção e hipnose. A banda alfa está no intervalo de 7,5-12,0 Hz, foi a primeira banda identificada por Hans Berger<sup>1</sup>, observada nas regiões posteriores da cabeça em ambos os hemisférios, região occipital, sendo maior em amplitude no hemisfério dominante (hemisfério responsável pelas habilidades mais desenvolvidas no indivíduo), também está associada ao relaxamento e aos processos não cognitivos. A banda mu está no intervalo de 8-12 Hz, é relacionada a realização e imaginação de movimentos, ela é observada na área sensorimotor do cérebro. A banda beta está no intervalo de 12-30 Hz, ocorre nos dois lados do cérebro de forma simétrica, pode ser dividida em beta 1 e beta 2 e tem atividade aumentada

<sup>1</sup> Hans Berger mostrou que é possível medir a atividade elétrica através do couro cabeludo (LUCK, 2014).

quando o indivíduo está resolvendo problemas matemáticos. A banda gama é a frequência acima de 30 Hz e reflete mecanismos de atenção, percepção e consciência (MOAZZAMI, 2012; RAMADAN *et al.*, 2015). Os componentes de interesse particular para BCI podem ser divididos em 4 categorias: atividade oscilatória do EEG, potenciais corticais lentos (*Slow Cortical Potentials* - SCP), potenciais neurais e potenciais relacionados a eventos (*Event-Related Potentials* - ERP) (VALLABHANENI; WANG; HE, 2005), este último será abordado na próxima seção.

As vantagens de utilizar o EEG é que os equipamentos de eletroencefalografia possuem custos menores que os equipamentos dos outros métodos, é portátil e com o auxílio de uma touca torna o posicionamento dos eletrodos mais simples. O EEG tem alta resolução temporal, capaz de medir as alterações na atividade cerebral em poucos milissegundos depois que ocorrem. Em relação a resolução espacial não é tão bom quanto os métodos invasivos. Devido a distância entre o local de manifestação e de onde é medida a atividade cerebral, o EEG apresenta uma baixa relação sinal ruído (*signal to noise ratio* - SNR), ou seja, há bastante interferência no sinal de EEG (HE *et al.*, 2013).

## 2.3 *Event-Related Potentials*

*Event-Related Potentials* (ERPs) são potenciais elétricos gerados pelo cérebro que são relacionados a eventos internos ou externos (estímulos, respostas, decisões). Eles são alterações no EEG que são *time-locked* a eventos, ou seja, são alterações que ocorrem após um evento que pode ser sensorial, cognitivo ou motor. Estes potenciais fornecem uma abordagem segura e não invasiva para estudar processos cognitivos (SUR; SINHA, 2009).

### 2.3.1 *Componentes do ERP*

Os ERPs consistem de uma sequência de componentes, estão presentes nas bandas de frequência anteriormente descritas e podem ser divididos em duas categorias principais: exógeno e endógeno. Os componentes exógenos são respostas obrigatoriamente a estímulos físicos e ocorrem devido ao processamento de eventos externos, mas independente do papel do estímulo no processamento da informação (VALLABHANENI; WANG; HE, 2005). As primeiras ondas, ou componentes, que ocorrem aproximadamente nos primeiros 100 ms após o início do estímulo, podem também ser nomeadas de sensoriais (SUR; SINHA, 2009).

Os componentes endógenos ocorrem quando um evento interno é processado. Ele é dependente do papel do estímulo na tarefa e da relação entre os estímulos e o contexto em que ocorreram. Estes ERPs refletem a forma em que o indivíduo avalia o estímulo enquanto examinam o processamento da informação (SUR; SINHA, 2009).

Um sinal luminoso aleatório irá gerar um componente exógeno, pois independentemente do contexto do sinal, o cérebro responde a sinais luminosos repentinos. Os componentes exógenos

ocorrem, por exemplo, quando um usuário está olhando para a letra que ele deseja em um soletrador e esta letra é apresentada (VALLABHANENI; WANG; HE, 2005).

Os componentes, também chamados de formas de onda, são descritos de acordo com sua latência e sua amplitude. Há componentes com picos positivos, tais como P50, P200 (ou P2) e P300 (ou P3) e componentes com picos negativos, tais como N100 (ou N1), N200 (ou N2) e N400. Cada um dos componentes possui uma latência diferente. Dentre os componentes do ERP, o P300, que é um pico positivo que ocorre aproximadamente 300 ms após apresentação do estímulo, é o que possui a maior amplitude (SUR; SINHA, 2009). Após apresentação de um estímulo, a presença do P300 é mais fácil de ser identificada no EEG que os demais componentes. Este é um dos motivos que o levaram a ser bastante utilizado em aplicações de BCIs, pois através da sua identificação é possível atingir um controle mais preciso da aplicação.

### 2.3.2 Componente P300

O P300, também chamado de P3, é um componente ERP exógeno que é desencadeado usando o paradigma *oddball* (RAMADAN *et al.*, 2015). Neste paradigma o usuário é apresentado a dois eventos, um raro (estímulo alvo) e outro frequente (estímulo não alvo). Quando o evento raro acontece, há um aumento na amplitude do sinal aproximadamente 300 ms após o início do evento. Acompanhado do evento raro, o usuário pode ser solicitado que conte quantas vezes esse evento acontece para aumentar a amplitude do P300 e facilitar a sua identificação. A amplitude do P300 é inversamente proporcional à frequência do evento raro, ela também reflete informações do estímulo, de modo que quanto maior atenção do indivíduo, maior é o P300 produzido. Normalmente é medido do córtex parietal (WOLPAW *et al.*, 2002). Farwell e Donchin (1988) utilizaram esse componente para desenvolverem o soletrador utilizando BCI. A principal vantagem do P300 é o alto número de opções que podem ser traduzidas com ele, sendo que apenas a opção desejada pelo usuário é que vai produzir o potencial P300.

A Figura 3 exemplifica como o soletrador baseado no P300 funciona. No lado direito representa as linhas e colunas que são intensificadas com o passar do tempo indicado no lado esquerdo. As letras *P* e *A* indicam se houve um estímulo padrão ou um estímulo alvo, respectivamente. Na figura, o caractere *B* está marcado com um círculo, indicando que o usuário está focando ele, ou seja, deseja selecionar este caractere. A primeira intensificação ocorre na terceira linha, por não conter o caractere desejado, é um estímulo padrão, então não produz o P300. A segunda intensificação ocorre na segunda coluna, o caractere desejado está nela, então é um estímulo alvo e por isso vai produzir o P300. Identificando a linha e coluna que ocorreram o P300 é possível identificar o caractere desejado pelo usuário.

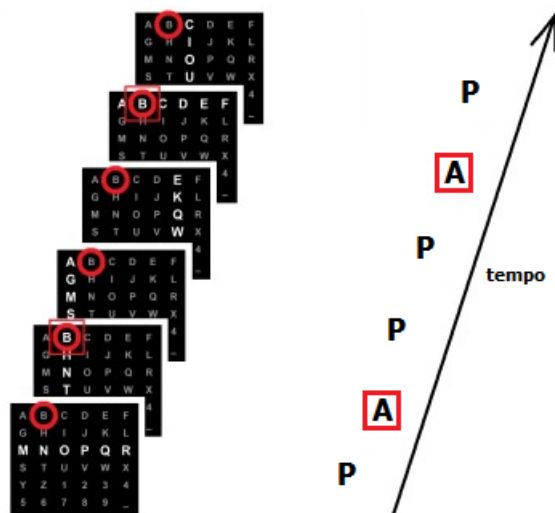


Figura 3 – Demonstração do paradigma *oddball*.

A Figura 4 apresenta um exemplo real de sinais de EEG com e sem o P300. Os sinais foram obtidos através da soma de sinais coletados no canal POz em um intervalo de tempo de aproximadamente 660 ms. Estes sinais também foram filtrados e tiveram sua taxa de amostragem diminuída. Na figura, o P300 encontra-se entre as amostras 6 e 8, correspondendo a aproximadamente 330 ms após a apresentação do estímulo.

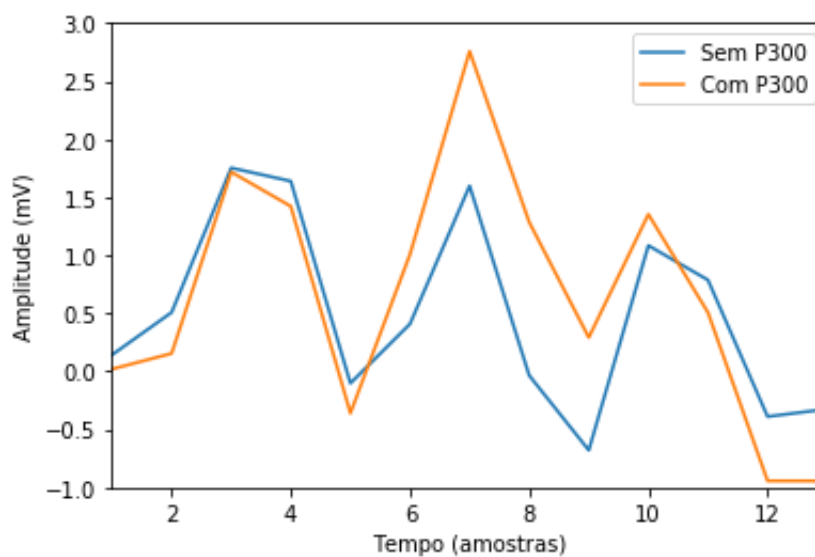


Figura 4 – Média dos sinais com e sem o P300.

## 2.4 Considerações Finais

Neste capítulo foram apresentados alguns conceitos introdutórios sobre BCI e sobre o componente P300, como funciona uma BCI de forma generalizada. Foi apresentada uma aplicação típica de BCI que é soletador baseado no P300.



---

# EXTRAÇÃO, SELEÇÃO E CLASSIFICAÇÃO

---

## 3.1 Considerações Iniciais

Neste capítulo são apresentados métodos para remoção de artefatos, filtros para filtrar os sinais, diminuição da taxa de amostragem dos sinais, também serão apresentadas formas para seleção dos canais e alguns algoritmos de aprendizado de máquina que são utilizados na classificação de sinais de EEG. Esta é a ordem das etapas que são realizadas no desenvolvimento de uma BCI.

## 3.2 Remoção de Artefatos e Ruídos

Os indivíduos são capazes de produzir voluntariamente sinais detectáveis para indicar suas intenções. Os sinais de EEG são coletados utilizando vários eletrodos espalhados no couro cabeludo. No entanto, também capturam sinais gerados de fontes fora do cérebro. Esses sinais são chamados de artefatos.

A presença deles no sinal de EEG compromete o funcionamento adequado da BCI. Desta forma, a remoção desse tipo de sinal é importante para que haja a maximização da relação sinal ruído e assim as características de interesse sejam extraídas dos sinais de EEG.

Para minimizar os efeitos dos ruídos é necessário saber a fonte deles. Os ruídos podem ser capturados através da atividade neural quando componentes não relacionados com os sinais desejados são capturados. Estes ruídos podem ser gerados por movimentos musculares, tais como, faciais e do pescoço, e afetam bastante os sinais desejados, prejudicando o desempenho dos classificadores. Este problema ainda é pior quando a amplitude e a frequência dos ruídos são semelhantes aos dos sinais da atividade cerebral ([NIJHOLT; TAN, 2010](#)).

Os artefatos, que não são gerados pelo sistema nervoso central, podem ser gerados pelo movimentos dos olhos (medido pelo Eletrooculograma - EOG), pelo batimento cardíaco (medido

pelo eletrocardiograma - ECG) e por outras fontes não neurais, são sinais mais proeminentes que os sinais de EEG. Uma das formas de tratar o ruído é solicitando para os indivíduos, que estão utilizando a BCI, para evitarem realizar movimentos com a cabeça durante os experimentos (VALLABHANENI; WANG; HE, 2005).

O descarte de sinais contaminados com ruídos e artefatos também pode ser realizado de forma manual, realizando uma análise e verificando que a amplitude é extremamente alta (BAŞAR-EROGLU *et al.*, 1992). Em vários experimentos com indivíduos sem problemas motores, é solicitado que eles se comportem como indivíduos que têm esclerose lateral amiotrófica, não são permitidos falar, mover os olhos e nem piscar, mas isso nem sempre funciona. Uma alternativa é separar os sinais de EOG e EMG e utiliza-los como mais uma fonte de informação de uma BCI (NIJHOLT; TAN, 2010). A remoção dos ruídos produzidos pelos movimentos do olho podem ser removidos através de métodos baseados no *blind source separation* (BSS) (JUNG *et al.*, 2000). A realização da gravação separada do EOG, com um eletrodo posicionado próximo ao olho, é utilizada para remoção de ruídos do EEG.

Além das interferências que são produzidas pelo movimento muscular, o EEG é vulnerável à interferência proveniente da rede elétrica. Mesmo a gravação sendo realizada dentro de um laboratório essa interferência ocorre, pois é difícil e caro realizar o isolamento dos aparelhos de medição do EEG. Geralmente o filtro Notch é aplicado à frequência de 50/60 Hz, permitindo a passagem de todas as frequências exceto as que estão na banda de rejeição (CORREIA *et al.*, 2010).

### 3.3 Extração de Características

Os sinais de EEG possuem várias características, que geralmente precisam ser processadas para serem obtidas. Cada uma delas representa uma propriedade diferente, uma medida reconhecível ou um componente funcional obtido de uma parte de um padrão. A etapa de extração de características tem como objetivo extrair informações ou características relevantes do sinal para que as características possam ser facilmente interpretadas.

As informações extraídas refletem a fisiologia e anatomia da atividade acontecendo dentro do cérebro. Há muitas variáveis em um grande conjunto de dados, o que requer uma grande quantidade de memória ou algoritmo poderoso para analisar os dados. Neste contexto, o método de extração de características é necessário para resolver essas variáveis ou informações a serem interpretadas de maneira simples e precisa (AL-FAHOUM; AL-FRAIHAT, 2014; AZLAN; LOW, 2014).

#### 3.3.1 Filtro Chebyshev Tipo I

Os filtros digitais têm como função remover partes do sinal não desejadas, como os ruídos, ou extrair partes de interesse, como frequências específicas. Há dois filtros digitais que



são bastante utilizados em sinais de EEG que são Chebyshev e Butterworth.

O filtro Chebyshev possui a característica de produzir *ripples*, que são ondulações nas bandas de frequências passantes, o número de *ripples* produzidos é igual a ordem do filtro. Este filtro possui alta precisão nas frequências de corte e atenuação acentuada (PERTENCE, 2015). Em Kusumandari, Hidayat e Turnip (2018) são comparados os filtros Chebyshev Tipo I e Butterworth, e os resultados obtidos mostram que o filtro Chebyshev consegue extrair melhor as frequências desejadas que o filtro Butterworth. Devido às características do filtro Chebyshev e resultados obtidos em trabalhos de BCI, ele foi escolhido para ser utilizado nos experimentos deste trabalho.

### 3.3.2 Normalização

Geralmente, a normalização de uma característica é necessária para neutralizar o efeito de diferentes características quantitativas que são medidas em diferentes escalas. Às vezes, a normalização de uma característica é realizada para restringir os valores da característica em uma escala, por exemplo entre 0 e 1 (FLACH, 2012). Há várias formas para normalizar uma característica em uma escala desejada. A função (3.1) pode ser utilizada para normalizar uma característica quantitativa entre 0 e 1, sendo  $x$  a característica.

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (3.1)$$

Outra forma de realizar a normalização é utilizando a função (3.2), chamada de *z-score*, onde  $\bar{x}$  é a média da característica e  $s$  é o desvio padrão. Utilizando esta função, a característica normalizada terá média 0 e variância unitária.

$$z_i = \frac{x_i - \bar{x}}{s} \quad (3.2)$$

Baseado em experimentos prévios e trabalhos relacionados (LI *et al.*, 2016; RAKOTOMAMONJY; GUIGUE, 2008), foi adotada a função (3.2) para normalização dos sinais de EEG utilizados neste trabalho.

### 3.3.3 Redução de Dimensionalidade

Após as etapas descritas anteriormente, os dados ainda não estão adequados para serem utilizados para o treinamento de uma BCI. No caso de sinais de EEG, a maioria dos trabalhos considera os segmentos dos sinais que correspondem aproximadamente a 600 ms após início do estímulo. Esses segmentos dos sinais que são utilizadas nos experimentos, pois são elas que possuem as informações relevantes (RAKOTOMAMONJY; GUIGUE, 2008; BOSTANOV, 2004; ABDULHAY *et al.*, 2017).

No entanto esses segmentos dos sinais possuem uma quantidade muito grande de dados. Por exemplo, na base utilizada neste trabalho (*Dataset II da BCI Competition III*) há 160 valores em aproximadamente 660 ms de um único canal. São utilizados 64 canais, ocorrem 180 estímulos para identificação de um único caractere e a base de treinamento possui 85 caracteres. É uma quantidade de dados muito grande que precisa ter sua dimensão diminuída. A redução da dimensionalidade conseqüentemente diminui a complexidade dos dados e melhora o desempenho dos algoritmos de aprendizado de máquina (AZLAN; LOW, 2014; JAIN; DUIN; MAO, 2000).

A Análise de Componentes Principais (*Principal Component Analysis - PCA*) é uma técnica que analisa dados compostos de variáveis dependentes que são, normalmente, interdependentes. É uma técnica utilizada para extrair informações importantes dos dados e expressar essa informação através de um novo conjunto de variáveis ortogonais chamadas componentes principais. Esta técnica também tem como objetivo diminuir o tamanho do conjunto de dados mantendo as informações importantes e por isso pode ser usada para extração de características e redução da dimensionalidade dos dados (XU *et al.*, 2004; ABDI; WILLIAMS, 2010).

Outra estratégia bastante utilizada para redução da dimensionalidade de bases constituídas por sinais de EEG é a decimação dos sinais. Partes específicas dos sinais passam por um processo de redução de amostragem por um fator. No presente trabalho, a base de dados foi construída a partir dos segmentos dos sinais que compreendem desde o início do estímulo até 667 ms após o início do estímulo. O fator para a redução da amostragem de cada um dos segmentos foi 12. Tanto o intervalo de 667 ms como o fator de amostragem 12 foram escolhidos baseados em outros trabalhos (RAKOTOMAMONJY; GUIGUE, 2008; ABDULHAY *et al.*, 2017). A PCA foi a primeira técnica para redução de dimensionalidade a ser testada neste trabalho e devido aos resultados iniciais insatisfatórios obtidos, preferiu-se adotar apenas a decimação.

### 3.4 Seleção de Canais

A redução de dados também é realizada através da seleção de canais. Este processo de selecionar canais reduz o número de eletrodos necessários para o funcionamento de uma BCI sem que ela diminua o desempenho, pois elimina os canais que possuem informações redundantes. O uso de canais desnecessários faz com que os algoritmos de classificação sofram com *overfitting* ao ruído (LAL *et al.*, 2004). Na literatura são mostrados alguns conjuntos de canais que podem ser utilizados e possibilitam o funcionamento de diversos tipos de BCI. Para as BCIs baseadas no P300 dois conjuntos de eletrodos são sugeridos: Fz, Cz, Pz, Oz, C3, C4, P3, P4, PO7 e PO8 (WANG; PENG, 2006; KRUSIENSKI *et al.*, 2008; ABDULHAY *et al.*, 2017; SELLERS; VAUGHAN; WOLPAW, 2010); PO7, Pz, CPz, P7, FC1, Cz, PO8 e FC5 (CASHERO, 2011).

A maior vantagem de utilizar canais pré-definidos é a capacidade de configurar rapidamente a BCI. No entanto, canais pré-definidos não garantem o melhor desempenho da BCI

para todos indivíduos e por isso diferentes estratégias são adotadas para selecionar um grupo de canais mais específico para cada indivíduo de forma que proporcione desempenhos melhores (KEE; PONNAMBALAM; LOO, 2015).

Os métodos de busca e otimização são uma das alternativas. Algoritmos Genéticos e Enxame de Partículas (*Particle Swarm Optimization* - PSO) são exemplos desses algoritmos que são utilizados para selecionar canais, apesar dos ótimos resultados que podem ser obtidos com eles, são métodos que têm o custo computacional muito grande (LAL *et al.*, 2004; MARTÍNEZ-CAGIGAL; SANTAMARÍA-VÁZQUEZ; HORNERO, 2019; KEE; PONNAMBALAM; LOO, 2015).

Os métodos de busca gulosa também são adotados com frequência. Assim como os algoritmos de busca e otimização, os métodos de busca gulosa também dependem do desempenho do classificador para avaliar o conjunto de canais selecionados. Eles funcionam de forma recursiva, eliminam os canais que menos contribuem ou adicionam canais que mais contribuem com o desempenho do classificador. A desvantagem desses métodos é que uma vez que o canal é eliminado do conjunto de canais ele não será mais utilizado nos testes seguintes, isso também ocorre quando o método adiciona um canal, todos os testes seguintes terão este canal e isto não garante que os canais selecionados ao final do processo sejam os melhores. A maior vantagem desses métodos é que possuem um custo computacional menor que os métodos de otimização e conseguem excelentes resultados (ARVANEH *et al.*, 2011; RAKOTOMAMONJY; GUIGUE, 2008).

Em Rakotomamonjy e Guigue (2008) é utilizado um algoritmo guloso que elimina canais. Neste algoritmo, o primeiro teste inicia com 64 canais selecionados e é avaliado o desempenho com esse conjunto de canais. Em seguida é removido temporariamente um canal e esse novo conjunto é avaliado, o canal que é removido temporariamente é escolhido de acordo com a ordem que está no arquivo, depois o canal retorna para o conjunto para outro canal ser temporariamente removido e o conjunto de canais ser novamente testado. Esse processo ocorre até que a remoção temporária de todos os canais termine. Depois são identificados os 4 canais que tiveram as melhores avaliações após sua eliminação e esses canais são removidos do conjunto por definitivo. Este processo continua até que fiquem apenas 4 canais. O processo de seleção de canais foi realizado para cada um dos modelos de um *ensemble*.

Nos resultados do trabalho anteriormente citado, foram selecionados no máximo, para um dos indivíduos do experimento, 16 canais e variando de 12 a 40 para o outro indivíduo. A seleção de canais tem como um dos objetivos melhorar o desempenho da BCI, então os canais que são selecionados são os canais que juntos atingiram o melhor desempenho dentre os conjuntos que foram testados. A quantidade de dados é bem menor que a dos 64 canais iniciais. Levando em consideração que a quantidade de dados que são manipulados começando com poucos canais e selecionando 16 é menor do que iniciando com 64 e selecionando 16, também considerando que a diminuição dos dados reduz a complexidade dos dados, facilitando o treinamento dos

classificadores e classificação, no presente trabalho foi adotado um algoritmo guloso para seleção de canais adotando uma estratégia de adição de canais para verificar possíveis diferenças nos resultados obtidos com o trabalho anteriormente citado.

O algoritmo guloso com adição de canais segue uma lógica semelhante ao descrito anteriormente, a estratégia adotada é apresentada no Algoritmo 1.

---

**Algoritmo 1** – Algoritmo de Seleção de Canais

---

```

canais ← [1..64]
canais_selecionados ← [ ]
avaliacoes_canais_selecionados ← [ ]
enquanto tamanho(canais_selecionados) < 60 faça
    avaliacoes = [ ]
    para todo canal em canais faça
        canais_selecionados.adicionar(canal)
        avaliacoes.adicionar(avaliar(canais_selecionados))
        canais_selecionados.remover(canal)
    fim para
    melhores_canais ← rank_canais(avaliacoes)
    canais_selecionados.adicionar(melhores_canais)
    temp ← avaliar(canais_selecionados)
    avaliacoes_canais_selecionados.adicionar(temp)
    canais.remover(melhores_canais)
fim enquanto

```

---

### 3.5 Classificadores

O funcionamento de uma BCI depende do desempenho de um classificador que foi treinado com sinais de EEG devidamente processados. O classificador após treinado é capaz de identificar a intenção do usuários a partir de novas entradas de sinais processados de EEG. Os classificadores para as BCI são treinados especificamente para um indivíduo e por isso o treinamento dos classificadores são realizados com sinais do indivíduo que utilizará a BCI.

Na literatura são mencionados diversos classificadores que são utilizados para a classificação de sinais de EEG. Dentre os mais utilizados são as Redes Neurais Artificiais (RNA), a Máquina de Vetores de Suporte (*Support Vectors Machine* - SVM) e a Análise Discriminante Linear (*Linear Discriminant Analysis* - LDA) (LOTTE *et al.*, 2007; AMIRI; FAZEL-REZAI; ASADPOUR, 2013).

O presente trabalho baseou-se no trabalho de Rakotomamonjy e Guigue (2008) para realizar as comparações de resultados, por isso a SVM foi escolhida. A LDA também foi utilizada para comparar o desempenho entre dois classificadores lineares, por ser um algoritmo bastante utilizado e obter ótimos resultados. A quantidade de parâmetros para serem refinados também contribuiu para a escolha desses algoritmos, a LDA não tem parâmetros para serem refinados

e a SVM possui um parâmetro. Visando a utilização do modelo computacional desenvolvido neste trabalho em uma aplicação *online*, a escolha dos algoritmos também considerou o tempo de execução de cada deles, pois o tempo é muito importante em aplicações de BCI *online*. Foram testadas Redes Neurais Artificiais variando número de neurônios e número de camadas. Foram testadas as SVMs, variando o *kernel* (polinomial, base radial e linear) e seus parâmetros. Também foi testada a LDA. A seguir serão apresentados os conceitos dos dois algoritmos que foram adotados neste trabalho, SVM linear e LDA.

A Máquina de Vetores de Suporte (*Support Vectors Machine* - SVM) é uma técnica de aprendizado de máquina desenvolvida originalmente para a classificação binária. A SVM realiza a classificação a partir de um hiperplano definido por um vetor  $w$  normal ao hiperplano e um termo "compensador"  $b$ . Tanto o vetor  $w$  como o termo  $b$  são determinados durante o treinamento da SVM. A Equação 3.3 define o hiperplano.

$$w \cdot x + b = 0 \quad (3.3)$$

Baseado nos dados de treinamento com  $l$  exemplos com vetores de características  $x_i$  e os rótulos das classes correspondentes  $y_i$ ,

$$(x_1, y_1), \dots, (x_l, y_l) \in \mathbb{R}^N \times \{-1, 1\} \quad (3.4)$$

a SVM precisa encontrar o hiperplano levando em consideração algumas restrições:

$$x_i \cdot w + b \geq +1 \quad \text{para } y_i = +1 \quad (3.5)$$

$$x_i \cdot w + b \leq -1 \quad \text{para } y_i = -1 \quad (3.6)$$

As Equações 3.5 e 3.6 podem ser combinadas em uma única equação:

$$y_i(x_i \cdot w + b) - 1 \geq 0 \quad \forall i \quad (3.7)$$

Seja  $x_1$  um ponto no hiperplano  $H_1 : w \cdot x + b = +1$  e  $x_2$  um ponto no hiperplano  $H_2 : w \cdot x + b = -1$ . Projetando  $x_1$  e  $x_2$  na direção de  $w$ , perpendicular ao hiperplano separador  $w \cdot x + b = 0$ , é possível ver a distância entre os hiperplanos  $H_1$  e  $H_2$ . Esta projeção é apresentada na Equação 3.8:

$$(x_1 - x_2) \left( \frac{w}{\|w\|} \cdot \frac{(x_1 - x_2)}{\|x_1 - x_2\|} \right) \quad (3.8)$$

A diferença entre  $w \cdot x_1 + b = +1$  e  $w \cdot x_2 + b = -1$  é igual a  $w \cdot (x_1 - x_2) = 2$ . Substituindo este resultado na Equação 3.8, tem-se:

$$\frac{2(x_1 - x_2)}{\|w\| \|x_1 - x_2\|} \quad (3.9)$$

Para obter o comprimento do vetor projetado, considera-se a norma da Equação 3.9 e tem-se:

$$\frac{2}{\|w\|} \quad (3.10)$$

Esta é a margem  $\gamma$ , apresentada na Figura 5, que é a distância entre os hiperplanos  $H_1$  e  $H_2$ , paralelos ao hiperplano ótimo. Maximizar a margem  $\gamma$  é, portanto, equivalente a minimizar  $(1/2)\|w\|^2$  com a restrição da Equação 3.7.

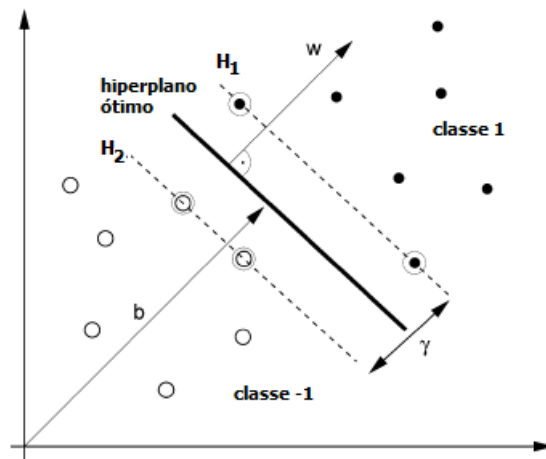


Figura 5 – Hiperplano da SVM para separar as classes maximizando a margem  $\gamma$ . Adaptado de (KAPER *et al.*, 2004).

Após a determinação do vetor  $w$  e do termo  $b$ , o rótulo da classe de novos dados  $x$  pode ser encontrado projetando  $x$  no vetor de pesos  $w$ , como apresentado na Equação 3.11.

$$f(x) = w \cdot x + b \quad (3.11)$$

Na Figura 5, não há exemplos dentro da margem  $\gamma$ , isso acontece porque as restrições impedem que isto ocorra. A SVM que foi descrita é com margens rígidas. Para permitir que essas restrições sejam violadas, são introduzidas as variáveis de folga tornando a margem suave. A introdução dessa variável permite a ocorrência de exemplos na margem  $\gamma$  e que exemplos sejam classificados incorretamente. Para maiores detalhes do algoritmo consultar as referências (BURGES, 1998; KAPER *et al.*, 2004; LORENA; CARVALHO, 2007; VAPNIK, 2013).

A Análise de Discriminantes Linear (*Linear Discriminant Analysis* - LDA), também conhecido como Discriminante Linear de Fisher, é um método estatístico multivariado que pode

ser utilizado para extração de características e classificação. O método pode ser utilizado se algumas condições forem satisfeitas, tal como: as classes de interesses devem ser mutuamente exclusivas e não podem ser perfeitamente correlacionadas. A LDA visa a criação de uma nova variável que é a combinação dos preditores originais. Isto é acompanhado pela maximização da diferença entre as classes em relação à nova variável.

Considerando  $X$  um vetor  $q$ -dimensional representando um exemplo de uma das possíveis classes. A LDA pode ser utilizada para classificar  $X$  se a classe é desconhecida. Um dos modos possíveis de descrever uma LDA é utilizando modelos probabilísticos. Supondo que a  $i$ -ésima classe tem densidade  $f_i(x)$  e probabilidade a priori  $\pi_i$ . A fórmula de Bayes diz:

$$P(\text{Classe} = i|x) = \frac{f_i(x)\pi_i}{\sum_j f_j(x)\pi_j} \quad (3.12)$$

A regra que classifica a maior probabilidade condicional faz com que seja menor o número esperado de erros de classificação. Isto é conhecido como Regras de Bayes ou Classificador de Bayes. Se incluirmos que a classe possui distribuição Gaussiana com média  $\mu_i$  e covariância  $\Sigma$ , então pode ser mostrada na classificação para a probabilidade condicional máxima é equivalente a classificar de acordo com a Equação 3.13,

$$\arg \max_i L_i, \quad (3.13)$$

onde  $L_i$  é a função discriminante e é mostrada na Equação 3.14.

$$L_i = x^T \Sigma^{-1} \mu_i - \frac{\mu_i^T \Sigma^{-1} \mu_i}{2} + \log \pi_i \quad (3.14)$$

No entanto, os parâmetros da distribuição Gaussiana não são conhecidos, e eles devem ser calculados através dos exemplos de treinamento de acordo com as Equações 3.15.

$$\begin{aligned} \hat{\pi}_k &= N_k/N, \text{ onde } N_k \text{ é o número de exemplos da classe } k; \\ \hat{\mu}_k &= \sum_{g_i=k} x_i / N_k; \\ \hat{\Sigma} &= \sum_{k=1}^K \sum_{g_i=k} (x_i - \hat{\mu}_k)(x_i - \hat{\mu}_k)^T / (N - K). \end{aligned} \quad (3.15)$$

A regra da LDA classifica para a classe 2 se satisfizer a condição da Equação 3.16, caso contrário a LDA classifica para a classe 1.

$$x^T \hat{\Sigma} (\hat{\mu}_2 - \hat{\mu}_1) > \frac{1}{2} \hat{\mu}_2^T \hat{\Sigma}^{-1} \hat{\mu}_2 - \frac{1}{2} \hat{\mu}_1^T \hat{\Sigma}^{-1} \hat{\mu}_1 + \log(N_1/N) - \log(N_2/N) \quad (3.16)$$

Para maiores detalhes sobre a LDA, consultar as referências (FRIEDMAN; HASTIE; TIBSHIRANI, 2001; THARWAT *et al.*, 2017; GARRETT *et al.*, 2003).

Foi mostrado neste capítulo as etapas realizadas para o desenvolvimento de uma BCI e o objetivo de cada etapa é de tornar o processo de classificação mais preciso. Uma das estratégias

utilizadas para atingir isto é a combinação de modelos computacionais, chamado de *ensemble* de modelos, esta técnica de aprendizado de máquina está entre as mais poderosas, pois diminui a influência da aleatoriedade de um único modelo.

O ponto principal é de como obter diversidade entre os diversos modelos. Uma das formas é utilizando diferentes conjuntos de treinamento para cada modelo, ou até mesmo diferentes características. Então, *ensemble* de métodos tem que seguir duas características em comum: (1) conter vários modelos preditivos e diversificados e (2) combinar a predição desses modelos de algum modo, geralmente através da média ou por voto.

Há dois *ensemble* de métodos mais conhecidos, *bagging* e *boosting*. *Bagging*, abreviação de *bootstrap aggregating*, é uma técnica que se caracteriza por criar vários modelos utilizando diferentes bases de dados criadas aleatoriamente a partir da base original. A probabilidade de escolher os exemplos é uniforme e com reposição, desta forma há probabilidade de exemplos não serem escolhidos assim como também de serem escolhidos mais de uma vez. É exatamente este o objetivo desta técnica, construir diferentes modelos a partir de diferentes bases de treinamento, desta forma que é inserida a diversidade entre as bases criadas.

O *boosting* utiliza uma estratégia um pouco diferente do *bagging*. Um algoritmo de aprendizado é executado várias vezes, em cada vez é fornecido uma base de treinamento com diferentes exemplos (com uma distribuição de exemplos diferente ou com exemplos ponderados). Cada vez que o algoritmo de aprendizado é executado ele gera uma nova regra de predição fraca e depois de várias rodadas, o algoritmo de *boosting* deve combinar essas regras fracas em uma única regra que será mais precisa que uma regra fraca. A distribuição de frequência e a ponderação dos exemplos aumenta à medida que o exemplo é classificado incorretamente. Isto força o algoritmo de aprendizado a dar mais atenção aos exemplos mais "difíceis" de acertar a classificação (FLACH, 2012; SCHAPIRE, 2003).

Em Bhatnagar *et al.* (2016) é apresentado um *ensemble* de SVMs com *kernel* de base radial para uma BCI baseada no P300. São utilizados 17 SVMs, a base de treinamento que é criada para cada uma das bases foi separada em 17 partições na ordem cronológica, não ocorrendo escolha aleatória dos exemplos para construção das bases. A classificação final realizada pelo *ensemble* é baseada na média da pontuação das SVMs. As características que são utilizadas é um vetor contendo a média dos dados das 12 linhas/colunas (6 linhas e 6 colunas) que foram intensificadas para um caractere. Nos experimentos realizados em Rakotomamonjy e Guigue (2008), também foi utilizado um *ensemble* de 17 SVMs, não utilizaram a média dos sinais e sim os próprios sinais como características, além de terem realizado uma seleção de canais, enquanto que no primeiro trabalho consideraram todos os canais. Essas são as principais diferenças entre os dois trabalhos.

No trabalho realizado em Kundu e Ari (2018) para detecção do P300, foi realizada uma comparação entre *ensemble* de SVMs e de LDAs. É utilizada um PCA para redução de dimensionalidade, de 10240 características para 680 e 619 (a base de dados utilizada continha



dados de dois indivíduos). Os resultados finais mostram que os dois métodos podem ser utilizados nas BCIs, com o *ensemble* de SVMs atingindo resultados melhores.

## 3.6 Considerações Finais

Neste capítulo foram detalhadas as etapas executadas no trabalho, destacando a importância de cada etapa, citando métodos alternativos que poderiam ser utilizados e os métodos que foram utilizados. Foi detalhado também o algoritmo de seleção de canais adotado no trabalho, também foram mencionados alguns trabalhos que serviram de base para o desenvolvimento deste trabalho.



---

## METODOLOGIA E RESULTADOS

---

### 4.1 Considerações Iniciais

Neste capítulo é apresentada a base de dados, os métodos utilizados para o pré-processamento e extração das características dos dados. Em seguida, são apresentados os resultados obtidos em diferentes testes. Em cada teste houve a variação ou de parâmetros ou da forma de seleção dos canais. Por fim, a comparação dos resultados obtidos com o de outros trabalhos que utilizam a mesma base de dados.

### 4.2 Descrição da Base de dados

A *BCI Competition* é uma competição de BCI. Em seu site são disponibilizadas bases de dados de EEG de vários laboratórios. Atualmente elas são utilizadas para pesquisas na área de BCI. A base de dados utilizada neste trabalho é o *Dataset II*<sup>1</sup> da *BCI Competition III*, competição realizada em 2004. Esta base contém a gravação do EEG de dois indivíduos sem problemas motores, usando um soletrador virtual, chamado também de *speller*. O EEG foi gravado utilizando o paradigma *oddball*, então há o potencial evocado P300 no EEG.

Para a coleta dos dados, os indivíduos foram apresentados a um teclado virtual que é uma matriz contendo 6 x 6 caracteres, com letras, números e o símbolo '-'. Cada um dos indivíduos tinha que focar um caractere específico no teclado e contar a quantidade de vezes que este caractere tinha sua cor intensificada. A cada sequência de 12 intensificações, uma em cada linha e em cada coluna, 2 das sequências produziram sinais com o P300.

Para cada caractere, foram coletados sinais referentes a 15 sequências de 12 intensificações (1 em cada linha e 1 em cada coluna), totalizando 180 intensificações para cada caractere, 30

---

<sup>1</sup> Mais detalhes sobre a base de dados podem ser encontrados na página da competição (<[http://bbci.de/competition/iii/desc\\_II.pdf](http://bbci.de/competition/iii/desc_II.pdf)>)

delas na linha ou coluna do caractere focado. Esta competição tinha como objetivo a identificação dos caracteres que cada um dos indivíduos estava focando considerando 5 e 15 sequências de 12 intensificações.

A matriz contendo os caracteres inicialmente fica 2,5 segundos sem nenhuma intensificação, durante este tempo é indicado ao usuário o caractere que deve ser focado. Em seguida as linhas e colunas, uma por vez, são intensificadas aleatoriamente no total de 15 sequências de 12 intensificações. Cada linha ou coluna é intensificada durante 100 ms e entre cada intensificação, a matriz fica sem nenhuma intensificação durante 75ms. Ao final das 15 sequências um outro caractere é apresentado por um tempo de 2,5 segundos e durante este tempo a matriz fica sem nenhuma intensificação. Todos os sinais coletados foram filtrados entre 0,1-60 Hz e digitalizados em 240 Hz.

Foi disponibilizada a base de treino e teste da gravação de EEG de 64 canais seguindo a distribuição 10-20. Na base de treino havia a gravação referente a 85 caracteres e na base de teste a gravação referente a 100 caracteres, para cada indivíduo. Os dados foram coletados em 5 sessões, não há informação de quantos caracteres foram coletados em cada sessão. Com as informações disponibilizadas na base de treino é possível identificar o momento de cada intensificação, a linha ou coluna que está sendo intensificada e se o indivíduo estava ou não focando a linha ou coluna que estava sendo intensificada. Na base de teste também há a identificação de qual linha e coluna estava sendo intensificada, mas não tem a informação de que o indivíduo estava ou não olhando a linha ou coluna que estava sendo intensificada. Como a competição já foi encerrada então esta informação pode ser encontrada no site da *BCI Competition*.

### 4.3 Pré-processamento dos Dados

Os dados dos sinais coletados, as informações das linhas e colunas que estavam sendo intensificadas em cada momento e se o indivíduo estava ou não focando a linha ou coluna intensificada estão em arquivos diferentes. Através do cruzamento das informações dos diferentes arquivos foi possível a identificação do momento exato do início de cada intensificação, que são os estímulos visuais necessários para a geração do P300. Desta forma os sinais foram segmentados considerando o intervalo de 667 ms após o início do estímulo. Foi considerado esse intervalo de tempo pois é um intervalo grande o suficiente para que ocorra o P300 (RAKOTOMAMONJY; GUIGUE, 2008; BHATNAGAR *et al.*, 2016).

A base de treino de cada indivíduo contém 85 caracteres, para cada caractere foram apresentadas 15 sequências de 12 intensificações, assim, no total para cada indivíduo há  $85 \times 15 \times 12 = 15300$  segmentos pós-estímulos, cada um dos sinais foi coletado de 64 canais. O intervalo de tempo de 667 ms equivale a 160 amostras de amplitude de um único canal. Esta separação foi realizada para cada um dos 64 canais, totalizando  $160 \times 64 = 10240$  características para cada segmento pós-estímulo. Cada um dos segmentos pós-estímulo foi rotulado com -1 e 1,

correspondendo a ausência e a presença do P300 respectivamente.

## 4.4 Seleção de Características

### 4.4.1 Filtragem

Os segmentos contendo 160 amostras dos sinais de cada canal foram filtrados utilizando o filtro Chebyshev I entre 0,1-10 Hz (RAKOTOMAMONJY; GUIGUE, 2008), 0,1-20 Hz (BHATNAGAR *et al.*, 2016; KUNDU; ARI, 2018; ABDULHAY *et al.*, 2017) e 0,1-30 Hz (DEEDWANIYA; GANDHI, 2016; KAPER *et al.*, 2004). Foram utilizados 3 intervalos de frequências diferentes porque há diferentes trabalhos que obtiveram bons resultados com cada um deles e realizando essa comparação será possível identificar qual é o mais adequado a ser utilizado para os experimentos deste trabalho. Os experimentos iniciais mostraram que a frequência de corte superior igual a 10 Hz atingiu melhores resultados, então utilizou-se apenas o intervalo de 0,1-10 Hz para executar todos os experimentos seguintes.

O filtro Chebyshev I possui um parâmetro referente a ondulação (*ripple*) nas bandas passantes. Este valor não é discutido nos trabalhos envolvendo P300, então 2 valores foram utilizados para comparar o desempenho com cada um deles e verificar qual o mais adequado para os experimentos. Os valores utilizados foram 0,1 dB e 0,5 dB, e a ordem do filtro foi 5. Foram construídas 6 bases (3 intervalos de frequência  $\times$  2 valores de ondulação) diferentes combinando os dois parâmetros.

### 4.4.2 Redução da dimensionalidade

Cada um dos segmentos do sinal contendo 160 valores de amplitude foi decimado pelo fator 12 resultando em 14 valores de amplitude ( $160/12 = 13,3$ , então aproxima-se para 14). Esta operação foi realizada em todas as instâncias, produzindo assim uma nova base contendo 15300 instâncias com  $14 \times 64$  canais = 896 características cada, 897 considerando o rótulo (-1 ou 1).

## 4.5 Ensemble de Classificadores

Nesta seção é detalhada a construção dos *ensembles* de classificadores. Baseando-se nos trabalhos discutidos e mencionados nos capítulos anteriores, inicialmente a base de dados foi dividida em 17 partições. A divisão foi realizada deixando os dados na sequência em que foram gravados, cada uma das partições possuía dados da gravação de 5 caracteres. Após a divisão, cada partição foi normalizada utilizando a Equação 3.2.

Em um primeiro experimento, foram construídos 2 *ensembles*, cada um deles contendo 17 modelos de um único algoritmo, os algoritmos utilizados foram SVM e LDA. A justificativa para escolha desses algoritmos foi apresentada no capítulo anterior.

Considerando  $K$  o índice da partição, se  $K \leq 9$ , então as partições de validação foram as *partição*<sub>1</sub> a *partição*<sub>9</sub> desconsiderando a partição  $K$ . Caso contrário,  $K > 9$ , então as partições de validação foram as *partição*<sub>10</sub> a *partição*<sub>17</sub> desconsiderando a partição  $K$ . Foi realizada essa divisão por considerar que há quantidade suficiente de dados para a validação (RAKOTO-MAMONJY; GUIGUE, 2008). O desempenho de cada modelo foi calculado de acordo com a Equação 4.1, também conhecida como Coeficiente de Jaccard. Onde  $tp$  são os verdadeiros positivos, são os exemplos classificados corretamente como contendo o P300,  $fp$  são os falsos positivos, são os exemplos classificados incorretamente como contendo o P300, e  $fn$  são os falsos negativos, os exemplo classificados incorretamente como não contendo o P300. O Coeficiente de Jaccard é utilizado para avaliar o desempenho da classificação de bases desbalanceadas e/ou quando apenas o desempenho na classificação de uma das classes que importa (RAMEZANI *et al.*, 2015). Neste trabalho, a base de dados é desbalanceada e as instâncias positivas são as que possuem o P300, são elas que importam, pois elas que irão ser utilizadas para controlar o soletrador. O valor deste coeficiente varia entre 0, quando não há nenhuma classificação correta, e 1, quando todas as classificações são corretas. Durante a fase de validação, as bases que foram utilizadas para validar o modelo foram normalizadas de acordo com os parâmetros calculados para normalizar a partição  $K$ , média e desvio padrão. Os canais que foram utilizados para o treinamento do modelo também são os mesmos que são considerados para validação e teste.

$$P_{validação} = \frac{tp}{tp + fp + fn} \quad (4.1)$$

A avaliação dos *ensembles* foi realizada utilizando a base de teste contendo dados de 100 caracteres para cada indivíduo. Os mesmos processos de filtragem, segmentação e decimação dos sinais foram realizados igualmente aos realizados na base de treinamento. Durante o teste, cada exemplo é classificado por todos os modelos do *ensemble*, cada modelo tem como saída a classe preditiva. Caso seja a classe positiva, é adicionado um voto na linha/coluna que o exemplo representa. Ao final da classificação de todos os exemplos de um caractere, é identificado linha/coluna que contém o maior número de votos e assim é realizada a identificação do caractere que o indivíduo estava focando.

O número de sequências corresponde à quantidade de vezes que as 6 linhas e as 6 colunas são intensificadas. Os resultados apresentados relacionados a 1 sequência significam que apenas os dados referentes a 1 sequência foram utilizados na avaliação, e assim por diante.

## 4.6 Experimentos e Resultados

Nesta seção são detalhados os experimentos realizados, os parâmetros testados e os resultados obtidos, assim como também a discussão dos mesmos. Os experimentos realizados seguiram as etapas anteriormente descritas. O primeiro experimento focou em refinar os parâmetros que poderiam ser utilizados em todos os outros experimentos.

As ordens do filtro Chebyshev Tipo I que foram testadas foram: 4, 5, 6 e 8. Esses valores foram utilizados baseando-se em outros trabalhos e de forma empírica. O filtro utilizado possui o parâmetro de ondulação nas bandas passantes, porém esse valor não é discutido nos trabalhos, então dois valores foram utilizados para identificar qual obteria os melhores resultados, 0,1 dB e 0,5 dB. A SVM possui o parâmetro de penalidade do erro, que são as variáveis de folga discutidas no capítulo anterior. Os valores testados foram: 0,01; 0,05; 0,1; 0,5 e 1,0 (RAKOTOMAMONJY; GUIGUE, 2008; KUNDU; ARI, 2018) e de acordo com os resultados obtidos foi adotado o valor 1,0 em todos os experimentos. As implementações dos algoritmos utilizados foram obtidas da biblioteca em python chamada Scikit-learn.

Os testes também compararam 2 algoritmos de aprendizado de máquina, SVM e LDA. Assim também como a quantidade de modelos utilizada no *ensemble*, 17 e 5 modelos, e também de forma isolada, com um único modelo. A seguir serão apresentados os resultados das combinações e as decisões feitas de acordo com os resultados. Os testes

#### 4.6.1 17 Classificadores

A combinação de cada um dos parâmetros discutidos anteriormente foi avaliada. A avaliação foi realizada de acordo com o descrito anteriormente, a ordem do filtro Chebyshev Tipo I igual a 5 e penalidade de erro igual a 1 apresentaram os melhores resultados, então esses valores serão utilizados em todos os outros experimentos. Nesses testes foram utilizados os 64 canais.

Nas Tabelas 1 e 2 são apresentados os desempenhos com a base filtrada com ondulação máxima 0,1 e 0,5 dB respectivamente. De 1 a 12 sequências a Tabela 2 apresenta o desempenho médio dos indivíduos ligeiramente melhor. Na literatura é relatado que a classificação do indivíduo A exige um maior número de canais (RAKOTOMAMONJY; GUIGUE, 2008). Na Tabela 1 o indivíduo A possui desempenho melhor que na Tabela 2 e também tem desempenho médio dos indivíduos igual ou melhor com 13, 14 e 15 sequências.

Tabela 1 – Número de caracteres classificados corretamente pelo *ensemble* de SVMs com 64 canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,1 dB.

Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	15	30	44	54	57	62	68	72	73	79	83	88	88	90	94
B	35	52	61	70	76	78	84	86	88	89	92	97	95	98	97
Média	25	41	52.5	62	66.5	70	76	79	80.5	84	87.5	92.5	91.5	94	95.5

Tabela 2 – Número de caracteres classificados corretamente pelo *ensemble* de SVMs com 64 canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,5 dB.

Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	14	30	46	58	59	61	68	72	78	82	86	92	88	92	92
B	35	57	63	66	80	81	80	86	92	92	95	95	95	94	96
Média	24,5	43,5	54,5	62	69,5	71	74	79	85	87	90,5	93,5	91,5	93	94

Nos resultados apresentados nas Tabelas 1 e 2, nenhuma apresentou os melhores desempenhos com os dois indivíduos e em todas as sequências, e por isso outros testes foram realizados. Como o presente trabalho visa avaliar a seleção de canais, então decidiu-se realizar testes com seleção de canais com os dois valores de ondulação máxima, 0.1 dB e 0,5 dB. Os resultados desses testes são apresentados a seguir.

Nas Tabelas 3 e 4 são apresentadas as avaliações de cada um dos modelos à medida que selecionam os canais. As células das tabelas que estão destacadas são as melhores avaliações de cada modelo. A avaliação foi realizada de acordo com a Equação 4.1. Comparando essas duas tabelas é possível identificar que a tabela do indivíduo A seleciona mais canais em cada um dos modelos, isso é relatado em Rakotomamonjy e Guigue (2008) e seleciona mais características como mostrado em Kundu e Ari (2018). O desempenho do indivíduo na classificação correta dos caracteres é menor que a do indivíduo B como apresentado nas Tabelas 1 e 2. Essa diferença ocorre devido às características dos sinais do próprio indivíduo, o estado mental durante as sessões de gravação pode influenciar bastante na qualidade dos sinais.

O número de caracteres classificados corretamente, pelos indivíduo A e B utilizando seleção de canais, é encontrado na Tabela 5. Novamente o indivíduo B obteve um desempenho melhor que o A, os resultados também foram melhores que os apresentados na Tabela 1, que utilizou os 64 canais nos 17 modelos. Os resultados não foram melhores exceto com 15 sequências.

A seleção de canais ocorre durante o treinamento dos modelos, cada um deles realizou sua própria seleção de canais. O algoritmo de seleção de canais utilizado neste trabalho pode ser encontrado no Algoritmo 1. Após a inclusão de novos canais, o novo conjunto com os canais selecionados é avaliado.



Tabela 3 – Avaliação de cada modelo para o indivíduo A com o aumento do número de canais através da seleção de canais com ondulação máxima 0,1 dB.

Partição	Número de canais utilizados									
	4	8	12	16	20	24	28	32	36	40
1	0,0	0,0	0,1322	0,1661	0,1802	0,1834	0,1890	0,1957	0,1922	0,2016
2	0,0	0,0	0,1360	0,1858	0,1950	0,2041	0,2045	0,1940	0,2090	0,2060
3	0,0	0,0	0,1377	0,2107	0,2181	0,2331	0,2243	0,2191	0,2204	0,2126
4	0,0	0,0	0,0000	0,0098	0,0687	0,1266	0,1618	0,1717	0,1680	0,1795
5	0,0	0,0	0,1731	0,1945	0,2181	0,2156	0,2119	0,2077	0,2130	0,1987
6	0,0	0,0	0,1637	0,2079	0,2192	0,2254	0,2195	0,2141	0,2232	0,2248
7	0,0	0,0	0,0072	0,1585	0,2030	0,2103	0,2109	0,2137	0,2177	0,2118
8	0,0	0,0	0,1485	0,2053	0,2064	0,2075	0,2268	0,2267	0,2252	0,2195
9	0,0	0,0	0,0065	0,1907	0,1891	0,2084	0,1992	0,1918	0,2012	0,2075
10	0,0	0,0	0,0000	0,1471	0,2130	0,2216	0,2290	0,2296	0,2169	0,2278
11	0,0	0,0	0,1314	0,1904	0,2181	0,2088	0,2161	0,2038	0,2119	0,2159
12	0,0	0,0	0,0914	0,1564	0,1982	0,2007	0,2098	0,2004	0,2089	0,2227
13	0,0	0,0	0,1097	0,1557	0,1822	0,2050	0,2086	0,2114	0,2099	0,2048
14	0,0	0,0	0,0813	0,1750	0,1875	0,2121	0,2228	0,2226	0,2087	0,2032
15	0,0	0,0	0,1725	0,2212	0,2218	0,2213	0,2204	0,2195	0,2196	0,2247
16	0,0	0,0	0,1496	0,1800	0,1990	0,1910	0,1922	0,1974	0,1990	0,2032
17	0,0	0,0	0,1252	0,1731	0,1999	0,2005	0,2014	0,1983	0,1966	0,2010

Tabela 4 – Avaliação de cada modelo para o indivíduo B com o aumento do número de canais através da seleção de canais com ondulação máxima 0,1 dB.

Partição	Número de Canais									
	4	8	12	16	20	24	28	32	36	40
1	0,0	0,0	0,1707	0,2382	0,2484	0,2528	0,2524	0,2496	0,2605	0,2577
2	0,0	0,0	0,00	0,1499	0,2245	0,2520	0,2596	0,2556	0,2439	0,2547
3	0,0	0,0	0,2242	0,2765	0,3256	0,3295	0,3177	0,3204	0,3117	0,2941
4	0,0	0,0	0,2247	0,2762	0,3098	0,3164	0,3161	0,3188	0,2977	0,3191
5	0,0	0,0	0,00	0,1876	0,2358	0,2609	0,2816	0,2620	0,2537	0,2606
6	0,0	0,0	0,1701	0,2684	0,3151	0,3184	0,3019	0,3088	0,2959	0,2946
7	0,0	0,0	0,2036	0,2739	0,2993	0,2918	0,2963	0,2898	0,2876	0,2929
8	0,0	0,1	0,2776	0,2864	0,2961	0,2941	0,2971	0,2997	0,2887	0,2854
9	0,0	0,0	0,2478	0,2873	0,3050	0,3064	0,2890	0,2907	0,2778	0,2866
10	0,0	0,0	0,2076	0,2928	0,2963	0,2869	0,2890	0,2897	0,2939	0,2831
11	0,0	0,0	0,1703	0,2603	0,2901	0,2908	0,2926	0,2888	0,2896	0,2797
12	0,0	0,0	0,1985	0,2770	0,3064	0,3176	0,3138	0,3110	0,3109	0,3052
13	0,0	0,0	0,1983	0,2768	0,2893	0,2977	0,3149	0,3138	0,2975	0,3039
14	0,0	0,0	0,1943	0,2874	0,2944	0,2887	0,2897	0,2967	0,2855	0,2921
15	0,0	0,0	0,1892	0,2955	0,3215	0,3128	0,3095	0,2975	0,2865	0,2811
16	0,0	0,0	0,0000	0,2691	0,2817	0,3023	0,3043	0,2911	0,2927	0,2762
17	0,0	0,0	0,2020	0,2738	0,2750	0,2703	0,2702	0,2788	0,2783	0,2544

Tabela 5 – Número de caracteres classificados corretamente pelo *ensemble* de SVMs após seleção de canais e aumentando o número de sequências utilizando dados dos indivíduos A e B filtrados com ondulação máxima de 0,1 dB.

Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	18	28	49	53	65	66	74	79	80	84	88	87	89	92	92
B	42	63	62	69	81	86	90	91	94	94	94	97	95	96	97
Média	30	45,5	55,5	61	73	76	82	85	86	89	91	92	92	94	94,5

A Tabela 6 mostra o desempenho obtido pelo indivíduo A e B com seleção de canais e com sinais filtrados com ondulação máxima de 0,5 dB. Quando comparado aos resultado da Tabela 5, que utiliza dados filtrados com ondulação máxima de 0,1 dB, pode-se perceber que não há superioridade absoluta nos resultados, dependendo do número de sequências uma tem melhor resultado que a outra, e às vezes iguais. Nas Tabelas 7 e 8 são apresentadas as avaliações de cada modelo, para os indivíduos A e B, com sinais filtrados com ondulação máxima 0,5 dB. As células das tabelas que estão destacadas são as melhores avaliações de cada modelo. A avaliação foi realizada de acordo com a Equação 4.1. É possível perceber que o indivíduo A seleciona mais canais que o indivíduo B, isso também foi encontrado nos resultados da Tabela 3.

Tabela 6 – Número de caracteres classificados corretamente pelo *ensemble* de SVMs após seleção de canais e aumentando o número de sequências e utilizando dados dos indivíduos A e B com ondulação máxima de 0,5 dB.

Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	18	30	45	50	63	65	75	79	82	80	86	89	91	90	92
B	47	57	64	73	81	87	89	90	93	92	92	96	94	95	96
Média	32,5	43,5	54,5	61,5	72	76	82	84,5	87,5	86	89	92,5	92,5	92,5	94

Comparando as Tabelas 3 e 7, são mostradas as avaliações do indivíduo A utilizando sinais filtrados com ondulação máxima 0,1 dB e 0,5 dB respectivamente, a quantidade de canais selecionados é menor com ondulação máxima 0,1 dB. A mesma comparação feita com o indivíduo B, analisando as Tabelas 4 e 8, a quantidade de canais é ligeiramente é menor com ondulação máxima 0,1 dB. Baseado nesses resultados, todos os outros testes foi utilizando o valor 0,1 dB para a ondulação máxima para o filtro Chebyshev.

Os resultados mais importantes que podem ser observado nas Tabelas 3, 4, 7 e 8 é que as avaliações, após selecionar 4 e 8 canais, continua 0. Esse comportamento foi inesperado. Utilizando a estratégia de inclusão de canais, inicialmente não há nenhum canal selecionado. Na primeira iteração do Algoritmo 1 as avaliações são de cada canal individualmente. A utilização de 1 único canal, independente dele, não conseguiu classificar nenhum caractere corretamente, então utilizando a Equação 4.1 para avaliar o desempenho do canal selecionado o resultado foi 0 e isso aconteceu para todos os 64 canais. A avaliação máxima que pode ser obtida é 1,

indicando que todas as instâncias foram classificadas corretamente. Então, quanto maior forem as avaliações dos modelos, maior será a capacidade do *ensemble* de identificar o caractere que o indivíduo está focando.

Diante desta situação, o algoritmo selecionou os canais que estavam na sequência disponibilizada no arquivo. Os 4 primeiros canais selecionados foram FC5, FC3, FC1 e FCz que não são os melhores para serem utilizados, principalmente se a quantidade de canais para serem usados for baixa. Na iteração seguinte, os canais que foram selecionados: FC2, FC4, FC6 e C5. Como apresentado nas tabelas, esses 8 canais não contribuíram para melhorar a classificação. Isso interferiu em todo o processo de seleção de canais porque a inclusão desses canais incluiu informações ruins no classificador, por serem canais bem próximos e não tanto relacionados com o P300 (WOLPAW *et al.*, 2002) podem ter inserido ruído ou informação redundante e consequentemente fez com que mais canais fossem selecionados.

Os próximos testes realizados utilizaram conjuntos de canais fixos que foram propostos na literatura (WANG; PENG, 2006; KRUSIENSKI *et al.*, 2008; ABDULHAY *et al.*, 2017; CASHERO, 2011). Na Tabela 9 são apresentados os resultados obtidos utilizando 2 conjuntos distintos para os indivíduos A e B. Lembrando que os testes realizados a partir deste ponto são utilizando sinais filtrados com ondulação máxima de 0,1 dB. Os resultados nesta tabela mostram, mais uma vez, que o indivíduo B obteve desempenho melhor que o indivíduo A nos dois conjuntos de canais. E o conjunto que possui mais canais, 10 canais, atingiu melhor resultado.

Tabela 7 – Avaliação dos classificadores para o indivíduo A com o aumento do número de canais através da seleção de canais com ondulação máxima de 0,5 dB.

Partição	Número de Canais									
	4	8	12	16	20	24	28	32	36	40
1	0,0	0,0	0,114	0,161	0,176	0,195	0,208	0,208	0,211	0,196
2	0,0	0,0	0,122	0,170	0,184	0,197	0,219	0,213	0,219	0,207
3	0,0	0,0	0,136	0,204	0,218	0,218	0,217	0,211	0,208	0,211
4	0,0	0,0	0,0	0,012	0,062	0,125	0,147	0,169	0,172	0,183
5	0,0	0,0	0,164	0,202	0,215	0,225	0,215	0,205	0,210	0,201
6	0,0	0,0	0,168	0,195	0,216	0,217	0,220	0,223	0,223	0,231
7	0,0	0,0	0,006	0,160	0,196	0,205	0,212	0,207	0,196	0,201
8	0,0	0,0	0,160	0,192	0,197	0,199	0,212	0,213	0,211	0,226
9	0,0	0,0	0,004	0,163	0,199	0,227	0,208	0,217	0,227	0,220
10	0,0	0,0	0,0	0,147	0,208	0,221	0,229	0,222	0,215	0,219
11	0,0	0,028	0,138	0,191	0,208	0,210	0,208	0,203	0,227	0,222
12	0,0	0,0	0,082	0,150	0,176	0,185	0,198	0,207	0,210	0,206
13	0,0	0,0	0,091	0,157	0,195	0,200	0,204	0,222	0,213	0,225
14	0,0	0,0	0,072	0,147	0,187	0,211	0,214	0,213	0,209	0,217
15	0,0	0,0	0,164	0,226	0,222	0,223	0,231	0,225	0,224	0,219
16	0,0	0,0	0,146	0,176	0,191	0,209	0,204	0,210	0,208	0,209
17	0,0	0,0	0,128	0,174	0,193	0,208	0,215	0,208	0,201	0,210

Tabela 8 – Avaliação dos classificadores para o indivíduo B com o aumento do número de canais através da seleção de canais com ondulação máxima de 0,5 dB.

Partição	Número de Canais									
	4	8	12	16	20	24	28	32	36	40
1	0,0	0,0	0,175	0,250	0,250	0,255	0,264	0,256	0,234	0,229
2	0,0	0,0	0,00	0,156	0,226	0,236	0,245	0,243	0,251	0,244
3	0,0	0,0	0,222	0,277	0,301	0,305	0,310	0,306	0,296	0,293
4	0,0	0,0	0,221	0,290	0,305	0,304	0,310	0,317	0,312	0,304
5	0,0	0,0	0,00	0,186	0,230	0,237	0,243	0,252	0,266	0,262
6	0,0	0,0	0,168	0,278	0,312	0,311	0,3165	0,315	0,297	0,299
7	0,0	0,0	0,210	0,266	0,305	0,314	0,304	0,297	0,290	0,294
8	0,0	0,161	0,276	0,308	0,320	0,304	0,307	0,302	0,286	0,2922
9	0,0	0,0	0,252	0,300	0,298	0,309	0,291	0,303	0,285	0,289
10	0,0	0,0	0,220	0,266	0,294	0,300	0,296	0,301	0,287	0,277
11	0,0	0,0	0,188	0,268	0,283	0,291	0,290	0,298	0,294	0,286
12	0,0	0,0	0,192	0,300	0,281	0,284	0,287	0,293	0,297	0,301
13	0,0	0,0	0,211	0,282	0,292	0,308	0,300	0,296	0,283	0,289
14	0,0	0,0	0,199	0,289	0,303	0,298	0,289	0,302	0,279	0,282
15	0,0	0,0	0,194	0,296	0,3077	0,309	0,317	0,305	0,302	0,305
16	0,0	0,0	0,00	0,254	0,296	0,303	0,295	0,273	0,281	0,269
17	0,0	0,0	0,206	0,273	0,287	0,271	0,285	0,279	0,270	0,265

Tabela 9 – Número de caracteres classificados corretamente pelo *ensemble* de 17 SVMs e aumentando o número de seqüências.

N. Sequência	Fz, Cz, Pz, Oz, C3, C4, P3, P4, PO7 e PO8														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Indivíduo A	12	15	25	31	34	41	43	44	52	56	56	59	63	68	71
Indivíduo B	30	43	57	63	73	81	83	85	87	87	88	92	94	95	95
Média	21	29	41	47	53,5	61	63	64,5	69,5	71,5	72	75,5	78,5	81,5	83
N. Sequência	PO7, Pz, CPz, P7, FC1, Cz, PO8 e FC5														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Indivíduo A	12	15	22	28	25	33	41	43	44	46	52	57	58	65	70
Indivíduo B	37	45	52	59	64	70	74	77	80	83	86	86	83	83	86
Média	24,5	30	37	43,5	44,5	51,5	57,5	60	62	64,5	69	71,5	70,5	74	78

Em Rakotomamonjy e Guigue (2008), o *ensemble* para o indivíduo A possui modelos que utilizavam apenas 4 canais. Baseado nisso e levando em consideração que a seleção de canais inicialmente sem nenhum conjunto de canal selecionado acabou selecionando 8 canais sem atingir bons resultados, então decidiu-se iniciar a seleção de canais com 4 canais fixos, sendo eles: Pz, Oz, PO7 e PO8. Este canais foram os selecionados porque de acordo com Rakotomamonjy e Guigue (2008), PO7 e PO8 foram os mais selecionados nos experimentos realizados para os dois indivíduos, Pz está presente nos dois conjuntos sugeridos na literatura e Oz é um dos canais mais relacionados com o P300.

A Tabela 10 apresenta o desempenho dos indivíduos A e B na classificação correta dos caracteres. Os modelos do *ensemble* realizaram a seleção de canais iniciando com os canais Pz, Oz, PO7 e PO8 pré-selecionados. O desempenho foi melhor que os apresentados anteriormente.

As Tabelas 11 e 12 apresentam os resultados da avaliação dos indivíduo A e B, respectivamente, com a seleção de canais iniciando com os canais fixos. Com essa alteração na seleção de canais o número de modelos com avaliação 0 do indivíduo A foi de 6, enquanto que do indivíduo B foi de apenas 1. Conseqüentemente houve uma diminuição nos canais selecionados do indivíduo B se comparada com a Tabela 4, enquanto que no indivíduo A não houve diferença na média de canais selecionados pelos modelos se comparado com os resultados apresentados na Tabela 3.

Tabela 10 – Número de caracteres classificados corretamente pelo *ensemble* de 17 SVMs após seleção de canais, iniciando com os canais Pz, Oz, PO7 e PO8 fixos, e aumentando o número de seqüências utilizando dados dos indivíduos A e B.

Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	15	34	54	52	64	66	76	73	75	83	87	91	91	92	95
B	41	59	72	79	81	86	87	90	92	94	93	97	94	96	97
Média	28	46,5	63	65,5	72,5	76	81,5	81,5	83,5	88,5	90	94	92,5	94	96

Tabela 11 – Avaliação dos *ensembles* com 17 SVMs para o indivíduo A com a seleção de canais iniciando com os canais Pz, Oz, PO7 e PO8 selecionados.

Partição	Número de Canais											
	4	8	12	16	20	24	28	32	36	40	44	
1	0,001	0,138	0,175	0,188	0,193	0,200	0,207	0,206	0,208	0,211	0,204	
2	0,079	0,183	0,200	0,204	0,208	0,200	0,195	0,203	0,204	0,210	0,204	
3	0,055	0,198	0,203	0,221	0,235	0,220	0,214	0,220	0,216	0,205	0,201	
4	0,0	0,0	0,011	0,082	0,125	0,150	0,175	0,193	0,190	0,194	0,202	
5	0,163	0,195	0,210	0,200	0,199	0,193	0,206	0,207	0,209	0,210	0,196	
6	0,163	0,204	0,206	0,215	0,219	0,225	0,225	0,226	0,213	0,220	0,210	
7	0,0	0,162	0,195	0,210	0,204	0,201	0,202	0,206	0,207	0,198	0,210	
8	0,002	0,152	0,172	0,183	0,185	0,194	0,193	0,206	0,215	0,220	0,231	
9	0,0	0,194	0,207	0,217	0,207	0,206	0,205	0,209	0,207	0,210	0,211	
10	0,0	0,102	0,204	0,221	0,218	0,224	0,209	0,210	0,216	0,219	0,211	
11	0,157	0,193	0,209	0,218	0,218	0,218	0,209	0,204	0,204	0,212	0,213	
12	0,0	0,143	0,204	0,221	0,224	0,208	0,217	0,215	0,213	0,209	0,208	
13	0,005	0,145	0,187	0,203	0,204	0,209	0,213	0,211	0,219	0,223	0,212	
14	0,0	0,106	0,172	0,192	0,204	0,211	0,221	0,215	0,203	0,199	0,203	
15	0,135	0,213	0,227	0,219	0,221	0,227	0,214	0,214	0,211	0,220	0,220	
16	0,102	0,164	0,198	0,203	0,203	0,202	0,204	0,193	0,204	0,207	0,207	
17	0,013	0,193	0,215	0,208	0,214	0,211	0,221	0,223	0,215	0,217	0,211	

Tabela 12 – Avaliação dos *ensembles* com 17 SVMs para o indivíduo B com a seleção de canais iniciando com os canais Pz, Oz, PO7 e PO8 selecionados.

Partição	Número de Canais										
	4	8	12	16	20	24	28	32	36	40	44
1	0,148	0,258	0,277	0,263	0,257	0,236	0,243	0,239	0,237	0,225	0,228
2	0,0	0,013	0,239	0,251	0,267	0,271	0,264	0,260	0,248	0,242	0,230
3	0,268	0,295	0,318	0,330	0,323	0,318	0,319	0,311	0,294	0,294	0,284
4	0,270	0,297	0,314	0,312	0,299	0,296	0,287	0,288	0,289	0,277	0,285
5	0,001	0,187	0,228	0,262	0,258	0,266	0,267	0,267	0,270	0,246	0,237
6	0,285	0,312	0,321	0,317	0,308	0,313	0,324	0,305	0,295	0,289	0,290
7	0,243	0,303	0,305	0,302	0,296	0,294	0,283	0,282	0,290	0,290	0,281
8	0,276	0,297	0,307	0,309	0,311	0,316	0,307	0,304	0,294	0,287	0,280
9	0,278	0,291	0,287	0,305	0,310	0,308	0,303	0,298	0,292	0,288	0,294
10	0,269	0,294	0,304	0,300	0,301	0,293	0,290	0,290	0,288	0,285	0,292
11	0,253	0,280	0,281	0,285	0,288	0,286	0,284	0,278	0,272	0,269	0,268
12	0,291	0,303	0,298	0,310	0,296	0,297	0,285	0,278	0,291	0,294	0,293
13	0,253	0,273	0,289	0,299	0,307	0,293	0,283	0,297	0,284	0,291	0,276
14	0,223	0,282	0,298	0,293	0,289	0,290	0,290	0,295	0,277	0,267	0,256
15	0,271	0,284	0,304	0,322	0,296	0,304	0,310	0,306	0,307	0,307	0,305
16	0,246	0,268	0,274	0,311	0,292	0,301	0,293	0,297	0,289	0,299	0,304
17	0,225	0,269	0,281	0,297	0,287	0,295	0,301	0,294	0,280	0,272	0,272

A quantidade de vezes que cada canal foi selecionado pelos modelos dos indivíduos A e B são apresentados nas Tabelas 14 e 15. Comparando com os resultados obtidos em (RAKOTOMAMONJY; GUIGUE, 2008), tanto a variedade como a quantidade de canais, selecionados com a seleção de canais utilizada no presente trabalho, foi maior. A seleção de canais conseguiu reduzir o número de canais e manteve ótimo desempenho para os dois indivíduos, porém esses resultados não foram melhores que os apresentados no trabalho citado anteriormente. Em Rakotomamonjy e Guigue (2008), o *ensemble* do indivíduo A selecionou 30 dos 64 canais, sendo que cada um desses canais foi selecionado por pelo menos 3 modelos e média de 29,88 canais selecionados por modelo. O *ensemble* do indivíduo B selecionou 28 dos 64 canais, sendo que cada um desses canais também foi selecionado por pelo menos 3 modelos e média de 9,29 canais selecionados por modelo. A diferença na quantidade de canais selecionados é maior no indivíduo B, comparando o presente trabalho com o anteriormente citado. No presente trabalho, os indivíduos A e B selecionaram 64 e 61 canais, respectivamente. Comparando o desempenho das duas abordagens, Tabelas 10 e 13, é possível perceber que com o número de sequências iguais a 1, 2, 3, 4 e 10, o desempenho do presente trabalho foi melhor, enquanto que com 5, 13 e 15 o desempenho foi inferior. Essa diferença é devido aos canais que foram selecionados.

Os experimentos utilizando todos os 64 canais, os dois conjuntos de canais propostos na literatura e a seleção de canais com os canais Pz, Oz, PO7 e PO8 pré-selecionados foram realizados também com um *ensemble* de LDAs. Os resultados podem ser conferidos na Tabela 16. Os resultados mostram que o *ensemble* de LDAs utilizando os 64 obteve um desempenho muito inferior ao *ensemble* de SVMs, isso mostra que a LDA tem dificuldade com dados com

muitas dimensões, e fica ainda mais nítido essa característica com a diminuição dos canais, o desempenho melhora bastante. Ao contrário do *ensemble* de SVMs, comparando os resultados com os dois conjuntos de canais fixos, o *ensemble* de LDAs obteve resultados melhores com o conjunto de 8 canais, enquanto o *ensemble* de SVMs obteve com o conjunto com 10 canais. O *ensemble* de LDAs atingiu resultados melhores com a seleção de canais, porém o *ensemble* de SVMs obteve melhor resultado que o *ensemble* de LDAs com a seleção de canais. Comparando com os resultados da Tabela 13, considerando com 1, 2, 3, 4 e 10 sequências, o *ensemble* de LDAs foi melhor.

Comparando as Tabelas 14 e 17, que mostram a quantidade de vezes que cada canal foi selecionado pelo *ensemble* do indivíduo A, é possível perceber que os canais mais e menos selecionados pelos dois *ensembles* são semelhantes, até o mesmo o fato de todos os canais terem sido selecionados pelo menos 1 vez por algum dos 17 modelos. A quantidade média de canais selecionados por cada um dos modelos é apresentado na Tabela 19, que mostra que o *ensemble* de LDAs selecionou menos canais para os dois indivíduos quando comparado com o *ensemble* de SVMs. A redução do número de canais é vantajosa por causa da diminuição de dados que menos colaboram para melhorar, ou até mesmo que prejudicam, o desempenho dos classificadores. Dessa forma torna o processo de aprendizado dos modelos mais eficiente (AZLAN; LOW, 2014).

Tabela 13 – Número de caracteres classificados corretamente por um *ensemble* de 17 SVMs (RAKOTO-MAMONJY; GUIGUE, 2008).

Indivíduos	Número de Sequências							
	1	2	3	4	5	10	13	15
A	16	32	52	60	72	83	94	97
B	35	54	62	68	75	91	96	96
Média	25,5	42,5	57	64	73,5	87	95	96,5

Tabela 14 – Quantidade de vezes que cada canal foi selecionado utilizando *ensemble* com 17 SVMs para o indivíduo A.

<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>
<b>17</b>	<i>Pz</i>	<b>10</b>	<i>Fz</i>	<b>9</b>	<i>PO3</i>	<b>6</b>	<i>CP1</i>
<b>17</b>	<i>Oz</i>	<b>10</b>	<i>FC6</i>	<b>8</b>	<i>FC5</i>	<b>6</b>	<i>F5</i>
<b>17</b>	<i>PO7</i>	<b>10</b>	<i>C1</i>	<b>8</b>	<i>FC2</i>	<b>6</b>	<i>FT8</i>
<b>17</b>	<i>PO8</i>	<b>10</b>	<i>C6</i>	<b>8</b>	<i>C3</i>	<b>5</b>	<i>T8</i>
<b>15</b>	<i>FC1</i>	<b>10</b>	<i>CP3</i>	<b>8</b>	<i>FP2</i>	<b>5</b>	<i>T9</i>
<b>13</b>	<i>FP1</i>	<b>10</b>	<i>CP4</i>	<b>8</b>	<i>AF3</i>	<b>5</b>	<i>T10</i>
<b>13</b>	<i>P7</i>	<b>10</b>	<i>F4</i>	<b>8</b>	<i>TP8</i>	<b>5</b>	<i>PO4</i>
<b>13</b>	<i>O1</i>	<b>10</b>	<i>F6</i>	<b>7</b>	<i>P3</i>	<b>4</b>	<i>C4</i>
<b>12</b>	<i>C5</i>	<b>9</b>	<i>FC4</i>	<b>7</b>	<i>Iz</i>	<b>4</b>	<i>F7</i>
<b>12</b>	<i>P6</i>	<b>9</b>	<i>CP2</i>	<b>7</b>	<i>FPz</i>	<b>4</b>	<i>F8</i>
<b>12</b>	<i>POz</i>	<b>9</b>	<i>CP6</i>	<b>7</b>	<i>F3</i>	<b>3</b>	<i>AF7</i>
<b>12</b>	<i>FC3</i>	<b>9</b>	<i>AFz</i>	<b>7</b>	<i>F2</i>	<b>3</b>	<i>AF4</i>
<b>11</b>	<i>FCz</i>	<b>9</b>	<i>F1</i>	<b>7</b>	<i>T7</i>	<b>3</b>	<i>AF8</i>
<b>11</b>	<i>Cz</i>	<b>9</b>	<i>FT7</i>	<b>7</b>	<i>P1</i>	<b>3</b>	<i>P4</i>
<b>11</b>	<i>CP5</i>	<b>9</b>	<i>TP7</i>	<b>6</b>	<i>P8</i>	<b>2</b>	<i>C2</i>
<b>11</b>	<i>CPz</i>	<b>9</b>	<i>P5</i>	<b>6</b>	<i>O2</i>	<b>2</b>	<i>P2</i>

Tabela 15 – Quantidade de vezes que cada canal foi selecionado utilizando *ensemble* com 17 SVMs para o indivíduo B.

<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>
<b>17</b>	<i>Pz</i>	<b>7</b>	<i>CP5</i>	<b>3</b>	<i>FC2</i>	<b>2</b>	<i>Fz</i>
<b>17</b>	<i>Oz</i>	<b>6</b>	<i>CP3</i>	<b>3</b>	<i>C6</i>	<b>2</b>	<i>F6</i>
<b>17</b>	<i>PO7</i>	<b>6</b>	<i>T7</i>	<b>3</b>	<i>AF3</i>	<b>2</b>	<i>FT8</i>
<b>17</b>	<i>PO8</i>	<b>5</b>	<i>TP8</i>	<b>3</b>	<i>AF4</i>	<b>2</b>	<i>T10</i>
<b>14</b>	<i>O1</i>	<b>5</b>	<i>PO4</i>	<b>3</b>	<i>AF8</i>	<b>2</b>	<i>P1</i>
<b>14</b>	<i>Iz</i>	<b>5</b>	<i>FC5</i>	<b>3</b>	<i>F7</i>	<b>2</b>	<i>P8</i>
<b>13</b>	<i>C3</i>	<b>5</b>	<i>CP4</i>	<b>3</b>	<i>F5</i>	<b>1</b>	<i>FC6</i>
<b>13</b>	<i>C5</i>	<b>5</b>	<i>FPz</i>	<b>3</b>	<i>F1</i>	<b>1</b>	<i>C2</i>
<b>13</b>	<i>CPz</i>	<b>4</b>	<i>F2</i>	<b>3</b>	<i>TP7</i>	<b>1</b>	<i>CP6</i>
<b>12</b>	<i>P2</i>	<b>4</b>	<i>F4</i>	<b>3</b>	<i>P6</i>	<b>1</b>	<i>F3</i>
<b>10</b>	<i>PO3</i>	<b>4</b>	<i>FT7</i>	<b>2</b>	<i>FC3</i>	<b>1</b>	<i>F8</i>
<b>10</b>	<i>POz</i>	<b>4</b>	<i>T9</i>	<b>2</b>	<i>C1</i>	<b>1</b>	<i>T8</i>
<b>9</b>	<i>Cz</i>	<b>4</b>	<i>P7</i>	<b>2</b>	<i>CP1</i>	<b>1</b>	<i>P5</i>
<b>9</b>	<i>FC1</i>	<b>4</b>	<i>P3</i>	<b>2</b>	<i>FP1</i>	<b>0</b>	<i>FC4</i>
<b>8</b>	<i>CP2</i>	<b>4</b>	<i>P4</i>	<b>2</b>	<i>AF7</i>	<b>0</b>	<i>C4</i>
<b>7</b>	<i>FCz</i>	<b>4</b>	<i>O2</i>	<b>2</b>	<i>AFz</i>	<b>0</b>	<i>FP2</i>



Tabela 16 – Número de caracteres classificados corretamente pelo *ensemble* de 17 LDAs utilizando diferentes combinações de canais.

64 Canais															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	11	8	10	14	19	16	16	21	25	25	30	32	36	32	34
B	4	12	15	15	16	19	22	26	26	34	36	39	41	38	40
Média	7,5	10	12,5	14,5	17,5	17,5	19	23,5	25,5	29,5	33	35,5	38,5	35	36
Fz,Cz,Pz, Oz, C3, C4, P3, P4, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	12	21	30	34	37	43	46	48	49	65	65	71	71	76	78
B	35	46	52	58	65	72	75	78	80	84	85	86	84	85	87
Média	23,5	33,5	41	46	56	57,5	60,5	63	64,5	74,5	75	78,5	77,5	80,5	82,5
PO7,Pz,CPz,P7, FC1,Cz, PO8 e FC5															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	14	20	23	34	41	42	50	53	58	67	67	67	68	74	78
B	35	51	57	68	74	81	83	87	89	89	89	92	89	92	93
Média	24,5	35,5	40	51	57,5	61,5	66,5	70	73,5	78	78	79,5	78,5	83	85,5
Seleção de Canais Iniciando com Pz, Oz, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	19	34	48	56	56	69	72	76	76	84	86	89	90	95	94
B	41	60	63	79	81	84	88	88	91	92	91	93	91	94	95
Média	30	47	55,5	67,5	68,5	76,5	80	82	83,5	88	88,5	91	90,5	94,5	94,5

Tabela 17 – Quantidade de vezes que cada canal foi selecionado utilizando 17 LDA's para o indivíduo A.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
17	Pz	8	AFz	5	P1	3	CP2
17	Oz	8	F6	5	P6	3	CP4
17	PO7	8	T7	4	FCz	3	FP2
17	PO8	7	C4	4	FC6	3	AF8
15	CPz	7	CP5	4	FPz	3	F5
12	O1	7	F1	4	AF4	3	Fz
12	Iz	6	Cz	4	F3	3	F8
10	FC4	6	C6	4	F2	3	TP7
10	T9	6	CP6	4	FT8	3	PO3
10	P7	6	F4	4	P3	3	PO4
10	P5	6	TP8	4	P2	2	FC3
9	POz	6	P4	4	P8	2	AF7
8	FC1	5	FC2	4	O2	2	F7
8	C5	5	CP3	3	FC5	2	T8
8	C3	5	FP1	3	C2	1	AF3
8	C1	5	T10	3	CP1	1	FT7

Tabela 18 – Quantidade de vezes que cada canal foi selecionado utilizando 17 LDA's para o indivíduo B.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
17	Oz	4	TP8	2	C1	1	AF7
17	PO7	4	P3	2	F5	1	AF3
17	PO8	4	PO3	2	F1	1	AF8
17	Pz	3	CP1	2	F6	1	F7
15	Iz	3	CP6	2	T7	1	F2
13	CPz	3	AFz	2	P7	1	F4
13	P2	3	F3	2	P1	1	T10
12	C3	3	F8	2	P4	1	P5
11	O1	3	FT7	2	PO4	1	O2
9	CP3	3	T8	1	FC2	0	FCz
6	Cz	3	T9	1	FC6	0	FC4
5	C5	3	TP7	1	C2	0	C4
5	CP2	3	P6	1	C6	0	FPz
5	POz	3	P8	1	CP4	0	AF4
4	FC1	2	FC5	1	FP1	0	Fz
4	CP5	2	FC3	1	FP2	0	FT8

Tabela 19 – Número médio de canais selecionados por cada um dos 17 modelos dos *ensembles* de SVMs e LDAs.

Algoritmo	Indivíduos	
	A	B
SVM	32,35	21,05
LDA	23,76	14,88

As Tabelas 20 e 21 apresentam resultados de outros trabalhos que utilizaram a mesma base de dados. A Tabela 20 apresenta resultados obtidos a partir da execução de *ensembles* ponderados, *Ensemble of Weighted Support Vector Machine* (EWSVM) e *Ensemble of Weighted Linear Discriminant Analysis* (EWLDA), a ponderação foi realizada levando em consideração avaliação calculada com a Equação 4.1. Neste trabalho também foi aplicado o PCA para redução de dimensionalidade. Mais detalhes podem ser consultados em Kundu e Ari (2018).

Os resultados apresentados na Tabela 21 são de dois trabalhos distintos que usaram um *ensemble* com 17 SVMs. Em Rakotomamonjy e Guigue (2008) é utilizada a seleção de canais com remoção de canais enquanto que em Bhatnagar *et al.* (2016) há a redução de exemplos através do cálculo da média dos sinais.

Os números de caracteres classificados corretamente pelos *ensembles* de 17 SVMs e 17 LDAs, Tabela 10 e 16 respectivamente, apresentaram desempenho médio melhor do que os modelos apresentados nas Tabelas 20 e 21, para os números de sequências de 1, 2, 3, 4, 5 e 10. Os algoritmos EWSVM e do apresentado em Bhatnagar *et al.* (2016) conseguiram melhor

desempenho com o indivíduo A do que o apresentado no presente trabalho considerando o número de sequências igual a 15.

Tabela 20 – Resultados obtidos por *ensembles* com modelos ponderados.

Algoritmo	Indivíduo	Número de Sequências							
		1	2	3	4	5	10	13	15
EWSVM	A	17	25	52	54	65	82	96	99
	B	39	62	67	75	79	93	92	97
	Média	28	43,5	59,5	64,5	72	87,5	94	98
EWLDA	A	16	23	30	42	47	76	84	86
	B	29	45	51	59	63	85	87	90
	Média	22,5	34	40,5	50,5	55	80,5	85,5	88

Tabela 21 – Resultados obtidos por *ensembles* de 17 SVMs.

Algoritmo	Indivíduo	Número de Sequências							
		1	2	3	4	5	10	13	15
Bhatnagar <i>et al.</i> (2016)	A	14	27	45	52	62	75	86	91
	B	36	50	61	73	79	89	93	94
	Média	25	38,5	53	62,5	70,5	82	89,5	92,5
Rakotomamonjy e Guigue (2008)	A	16	32	52	60	72	83	94	97
	B	35	53	62	68	75	91	96	96
	Média	25,5	42,5	57	64	73,5	87	95	95,5

#### 4.6.2 5 Classificadores

Foram realizados experimentos com *ensembles* de SVMs e LDA, cada um contendo 5 modelos. De forma semelhante como foi realizado com 17 modelos, foi realizada a divisão da base em 5 partições, cada partição possui sinais de 17 caracteres consecutivos. Foram considerados todos os sinais de todos os caracteres. Após a divisão da base, cada uma das bases foi normalizada. As mesmas etapas de processamento realizadas com os *ensemble* de 17 modelos também foram realizadas para os de 5, a única diferença é a quantidade de modelos.

É de conhecimento da área de BCI que os sinais são diferentes entre indivíduos e também há diferenças entre os sinais do mesmo indivíduo em resposta a um mesmo estímulo. Levando isso em consideração foram realizados experimentos com a base de dados dividida em 5 partições; este valor é devido aos dados terem sido coletados em 5 sessões, com isso todos os dados da mesma sessão estariam juntos. Enquanto que com 17 partições, há partições que contêm sinais que foram coletados em diferente sessões. Agrupando os sinais da mesma sessão há uma maior semelhança entre os sinais, considerando que as condições do ambiente se mantiveram constantes (BHATNAGAR *et al.*, 2016).

Na Tabela 22 são apresentados os resultados utilizando um *ensemble* com 5 SVMs e com diferentes conjuntos de canais. Comparando com as Tabelas 1 e 9 em relação à combinação

de canais equivalente, o *ensemble* com 17 modelos foi melhor em todos. Utilizando o conjunto com 10 canais fixos foi melhor que com 8 canais fixos. O indivíduo B obteve desempenho melhor que o do A, apenas com 64 canais que o desempenho de A não foi tão inferior. Ocorre também instabilidade no desempenho, a tendência era de que com o aumento do número de sequências houvesse a melhoria no desempenho, porém várias vezes o desempenho diminui e depois aumenta. Isso corre por exemplo, no indivíduo B com 10 canais fixos e com 13 sequências, também ocorre com indivíduo A com 8 canais e com 6, 9 e 13 sequências.

A diminuição e a instabilidade do desempenho podem estar associadas à maneira como é realizado o voto, como não há uma ponderação no voto, então isso permite que ocorra empate. E com a diminuição da quantidade de modelos a possibilidade de empate aumenta. Isso foi encontrado através da análise do número de votos. Outra observação importante foi o tempo de execução dos modelos, com 17 foi mais rápido que com 5. Isso ocorreu pelo aumento da dificuldade de encontrar um hiperplano devido ao aumento de dados, o que tornou a tarefa mais complexa (VAPNIK, 2013). Uma alternativa a este problema é diminuir o número de exemplos de amostra, com 17 modelos de SVMs foram utilizados 900 exemplos de treinamento em cada amostra e mostrou-se suficiente. Outra forma de diminuir a quantidade de exemplos de treinamento é realizando a média de sinais como realizado em Bhatnagar *et al.* (2016).

Tabela 22 – Número de caracteres classificados corretamente pelo *ensemble* de 5 SVMs com diferentes combinações de canais e aumentando o número de sequências.

64 Canais															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	20	30	37	40	47	57	59	71	70	77	81	86	87	89	93
B	39	55	59	65	74	82	86	87	90	90	91	93	93	95	94
Média	29,5	42,5	48	52,5	60,5	69,5	72,5	79	80	83,5	86	89,5	90	92	93,5
Fz,Cz,Pz, Oz, C3, C4, P3, P4, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	13	17	24	22	24	25	29	36	40	46	51	51	55	56	62
B	33	42	58	61	68	72	79	82	80	88	85	88	87	93	93
Média	23	29,5	41	41,5	46	48,5	54	59	60	67	68	69,5	71	74,5	77,5
PO7,Pz,CPz,P7, FC1,Cz, PO8 e FC5															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	7	12	12	15	20	18	21	22	21	28	30	30	29	31	35
B	29	36	53	58	61	64	69	73	72	79	79	83	80	81	83
Média	18	24	32,5	36,5	40,5	41	45	47,5	46,5	53,5	54,5	56,5	54,5	56	59
Seleção de Canais Iniciando com Pz, Oz, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	16	26	35	41	50	47	51	62	69	72	75	83	80	85	87
B	35	45	59	69	72	77	81	81	90	92	92	94	96	93	93
Média	25,5	35,5	47	55	61	62	66	71,5	79,5	82	83,5	88,5	88	89	90

Tabela 23 – Quantidade de vezes que cada canal foi selecionado utilizando 5 SVMs para o indivíduo A.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
5	FC5	5	Oz	3	CP3	3	P7
5	FC3	5	Pz	3	CPz	3	P1
5	FC1	4	C1	3	CP2	3	PO3
5	FCz	4	CP5	3	FP1	3	O1
5	FC2	4	CP4	3	FP2	3	O2
5	FC4	4	CP6	3	AF7	2	CP1
5	FC6	4	F3	3	F7	2	AF3
5	C5	4	F1	3	Fz	2	AFz
5	Cz	4	F2	3	F4	2	AF8
5	C2	4	F6	3	F8	2	F5
5	C6	4	T8	3	FT7	2	T10
5	FPz	4	C3	3	FT8	2	P5
5	AF4	4	P4	3	T7	2	P3
5	PO7	4	PO4	3	T9	2	P2
5	POz	4	Iz	3	TP7	2	P6
5	PO8	3	C4	3	TP8	2	P8

Tabela 24 – Quantidade de vezes que cada canal foi selecionado utilizando 5 SVMs para o indivíduo B.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
5	PO7	3	C6	2	C4	2	POz
5	PO8	3	CP6	2	CP3	2	Iz
5	Pz	3	FPz	2	CPz	1	CP5
5	Oz	3	Fz	2	CP2	1	FP1
4	FC5	3	F2	2	CP4	1	AF3
4	FC3	3	F8	2	FP2	1	AFz
4	FC1	3	T10	2	AF4	1	F4
4	FCz	3	TP8	2	AF8	1	FT7
4	FC2	3	P2	2	F7	1	FT8
4	FC4	3	PO4	2	F5	1	TP7
4	CP1	3	O1	2	F3	1	P7
4	P1	3	Cz	2	F1	1	P3
4	P6	3	O2	2	T7	1	P4
4	P8	2	C3	2	T8	0	AF7
3	FC6	2	C1	2	T9	0	F6
3	C5	2	C2	2	PO3	0	P5

A quantidade de vezes que cada canal foi selecionado pelos 5 modelos do *ensemble* de SVMs é apresentada nas Tabelas 23 e 24. Como já discutido anteriormente, o *ensemble* do indivíduo A selecionou mais canais que o do B e todos os canais foram selecionados, sendo que pelo menos por 3 modelos. Enquanto que o *ensemble* do indivíduo B 3 canais não foram selecionados por nenhum modelo. Alguns canais que são bastante relacionados com o P300 não foram selecionados, como por exemplo o canal CPz no indivíduo B, enquanto que no indivíduo

A ele foi selecionado 3 modelos. Esse canal está entre os 12 canais mais selecionados em Rakotomamonjy e Guigue (2008), no entanto eles realizaram testes com 17 modelos de SVM. Com a diminuição nos dados de treinamento provavelmente esses resultados seriam diferentes e um número menor de canais seriam selecionados.

Os resultados obtidos pelo *ensemble* com 5 LDAs utilizando diferentes combinações de canais são apresentados na Tabela 25. Os resultados com 64 e 10 canais são semelhantes, até 6 sequências o *ensemble* com 10 canais atingiu desempenho médio melhor, e a partir de 7 sequências o *ensemble* com 64 canais obteve melhor desempenho. Comparando com as demais combinações de canais, a que utiliza 8 canais obteve o pior desempenho. A seleção de canais permitiu atingir o melhor resultado comparado com os outros resultados apresentados na tabela.

Comparando os *ensembles* de 17 (Tabela 16) e 5 LDAs, houve um aumento no desempenho com 64 canais. Com 10 canais, os resultados foram semelhantes, porém até 6 sequências o desempenho foi melhor com 17 modelos. Utilizando 8 canais e com seleção de canais, houve uma diminuição no desempenho, sendo mais perceptível no indivíduo A.

Tabela 25 – Número de caracteres classificados corretamente pelo *ensemble* de 5 LDAs utilizando diferentes combinações de canais.

64 Canais															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	10	20	26	28	38	40	49	56	58	65	71	72	70	77	79
B	20	27	40	49	57	65	64	71	77	79	78	82	84	86	88
Média	15	23,5	33	38,5	47,5	52,5	56,5	63,5	67,5	72	74,5	77	77	81,5	83,5
Fz,Cz,Pz, Oz, C3, C4, P3, P4, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	11	19	26	30	32	36	41	52	55	58	60	65	65	72	77
B	31	36	51	60	70	73	79	79	80	85	84	87	89	90	88
Média	21	27,5	38,5	45	51	54,5	60	65,5	67,5	71,5	72	76	77	81	82,5
PO7,Pz,CPz,P7, FC1,Cz, PO8 e FC5															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	9	17	20	19	27	31	31	38	41	50	52	56	56	60	66
B	32	37	50	52	61	66	68	69	73	78	79	83	83	85	90
Média	20,5	27	35	35,5	44	48,5	49,5	53,5	57	64	65,5	69,5	69,5	72,5	78
Seleção de Canais Iniciando com Pz, Oz, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	9	20	28	35	47	52	56	56	69	75	76	78	82	86	90
B	39	52	65	71	73	79	84	85	90	92	93	92	92	91	94
Média	24	36	46,5	53	60	65,5	70	70,5	79,5	83,5	84,5	85	87	88,5	92

Tabela 26 – Quantidade de vezes que cada canal foi selecionado com 5 LDAs para o indivíduo A.

<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>	<b>Quant.</b>	<i>Canais</i>
<b>5</b>	<i>CPz</i>	<b>4</b>	<i>P7</i>	<b>3</b>	<i>Iz</i>	<b>2</b>	<i>PO4</i>
<b>5</b>	<i>T7</i>	<b>3</b>	<i>FC1</i>	<b>2</b>	<i>FCz</i>	<b>1</b>	<i>FC2</i>
<b>5</b>	<i>T8</i>	<b>3</b>	<i>Cz</i>	<b>2</b>	<i>FC4</i>	<b>1</b>	<i>FC6</i>
<b>5</b>	<i>T9</i>	<b>3</b>	<i>C4</i>	<b>2</b>	<i>C3</i>	<b>1</b>	<i>C2</i>
<b>5</b>	<i>Pz</i>	<b>3</b>	<i>CP1</i>	<b>2</b>	<i>C6</i>	<b>1</b>	<i>F5</i>
<b>5</b>	<i>PO7</i>	<b>3</b>	<i>CP6</i>	<b>2</b>	<i>CP5</i>	<b>1</b>	<i>Fz</i>
<b>5</b>	<i>POz</i>	<b>3</b>	<i>FP1</i>	<b>2</b>	<i>CP3</i>	<b>1</b>	<i>F2</i>
<b>5</b>	<i>PO8</i>	<b>3</b>	<i>FPz</i>	<b>2</b>	<i>CP4</i>	<b>1</b>	<i>F8</i>
<b>5</b>	<i>Oz</i>	<b>3</b>	<i>FP2</i>	<b>2</b>	<i>AF7</i>	<b>1</b>	<i>FT8</i>
<b>4</b>	<i>C5</i>	<b>3</b>	<i>AF8</i>	<b>2</b>	<i>F3</i>	<b>1</b>	<i>TP7</i>
<b>4</b>	<i>C1</i>	<b>3</b>	<i>F1</i>	<b>2</b>	<i>F4</i>	<b>1</b>	<i>P8</i>
<b>4</b>	<i>CP2</i>	<b>3</b>	<i>F6</i>	<b>2</b>	<i>P3</i>	<b>1</b>	<i>O2</i>
<b>4</b>	<i>AF3</i>	<b>3</b>	<i>T10</i>	<b>2</b>	<i>P1</i>	<b>0</b>	<i>FC5</i>
<b>4</b>	<i>AFz</i>	<b>3</b>	<i>TP8</i>	<b>2</b>	<i>P2</i>	<b>0</b>	<i>FC3</i>
<b>4</b>	<i>AF4</i>	<b>3</b>	<i>P5</i>	<b>2</b>	<i>P4</i>	<b>0</b>	<i>F7</i>
<b>4</b>	<i>FT7</i>	<b>3</b>	<i>O1</i>	<b>2</b>	<i>P6</i>	<b>0</b>	<i>PO3</i>

Fazendo um comparativo entre os *ensembles* com 5 LDAs e 5 SVMs é possível verificar que com 64 canais o *ensemble* de SVMs atingiu resultados melhores, mostrando que é mais robusto que o *ensemble* de LDAs com dados de grande dimensionalidade. Comparando os resultados obtidos pelos dois *ensembles* utilizando 10 e 8 canais, a média do desempenho dos modelos de LDAs se sobressaíram aos de SVMs e a diferença de desempenho é maior no indivíduo A. A seleção de canais produziu resultados semelhantes no desempenho médio entre os dois *ensembles*.

O *ensemble* de 5 LDAs para o indivíduo B selecionou menos canais que o do A, alguns canais não foram selecionados por nenhum modelo, isso foi observado anteriormente. Na Tabela 28 são apresentadas as médias de canais selecionados pelos modelos dos *ensembles*. E assim como apresentado na Tabela 19, novamente o *ensemble* de LDAs selecionou menos canais. O aumento de canais diminui o desempenho da LDA, enquanto que a SVM não sofre muito com o aumento de canais (aumento da dimensionalidade).

Tabela 27 – Quantidade de vezes que cada canal foi selecionado com 5 LDAs para o indivíduo B.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
5	PO7	2	C3	1	FPz	1	P5
5	POz	2	C2	1	FP2	1	P6
5	PO8	2	C6	1	AFz	1	PO3
5	Pz	2	CP3	1	AF4	0	FC5
5	O1	2	CP1	1	F7	0	FC3
5	Oz	2	CP2	1	F5	0	FC1
4	CPz	2	F4	1	F3	0	C4
4	Cz	2	TP8	1	Fz	0	CP5
4	P2	2	P7	1	F2	0	CP4
4	Iz	2	P3	1	F8	0	AF7
3	FCz	2	P1	1	FT7	0	AF3
3	FC4	2	PO4	1	FT8	0	AF8
3	C5	1	FC2	1	T7	0	F1
3	C1	1	FC6	1	T8	0	F6
3	T9	1	CP6	1	T10	0	P4
3	P8	1	FP1	1	TP7	0	O2

Tabela 28 – Número médio de canais selecionados por cada um dos 5 modelos dos *ensembles* de SVMs e LDAs.

Algoritmo	Indivíduos	
	A	B
SVM	13,52	9,29
LDA	9,64	6,53

### 4.6.3 1 Classificador

Os experimentos realizados com um único classificador foram diferentes dos executados anteriormente, cuja diferença encontra-se na divisão da base. Para um único classificador, a base foi dividida utilizando o método *hold-out*, que consiste em dividir a base de dados em dois grupos mutuamente exclusivos, um para treinamento e outro para teste. A divisão levou em conta os dados por caracteres. De forma aleatória foram escolhidos 60 caracteres para treinamento e 25 para teste. Como o treinamento da SVM demorou muito tempo, esse experimento foi executado apenas uma vez. A redução da base de treinamento reduziria o tempo de execução.

Os resultados dos testes realizados com uma única SVM são apresentados na Tabela 29. Os resultados com uma única SVM não foram os melhores, o desempenho com qualquer combinação de canais foi inferior a todos os outros apresentados nas Tabelas 9 e 22. Utilizando os conjuntos com 10 e 8 canais, a SVM para o indivíduo A foi incapaz de realizar a classificação corretamente.

A seleção de canais da SVM para o indivíduo A selecionou todos os 64 canais, enquanto que para o indivíduo B selecionou 40 canais. Mostrando mais uma vez que o indivíduo A precisa



de mais canais para conseguir um bom desempenho. Os canais selecionados para o indivíduo B são apresentados na Tabela 30.

Tabela 29 – Número de caracteres classificados corretamente por uma única SVM utilizando diferentes combinações de canais.

64 Canais															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	11	13	18	22	29	32	39	47	46	52	53	59	57	64	66
B	25	31	44	55	57	62	68	68	73	75	77	79	82	81	82
Média	18	22	31	38,5	43	47	53,5	57,5	59,5	63,5	65	69	69,5	72,5	74
Fz,Cz,Pz, Oz, C3, C4, P3, P4, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
B	21	33	39	41	45	54	60	66	70	73	73	76	76	77	80
PO7,Pz,CPz,P7, FC1,Cz, PO8 e FC5															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
B	24	24	34	43	49	50	59	63	61	65	64	68	69	72	71
Seleção de Canais Iniciando com Pz, Oz, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	11	13	18	22	29	32	39	47	46	52	53	59	57	64	66
B	29	41	57	63	66	72	72	79	85	85	84	87	87	88	90
Média	20	27	37,5	42,5	47,5	52	55,5	63	65,5	68,5	68,5	73	72	76	78

A LDA com 64 canais atingiu o melhor resultado como mostrado na Tabela 31, e se comparado com o *ensembles* de 17 LDAs utilizando 64 canais também foi melhor. Superou até mesmo a LDA com seleção de canais. O pior desempenhos foi com o indivíduo A utilizando 8 e 10 canais. Pelo menos a LDA conseguiu realizar a classificação com 10 e 8 canais para o indivíduo A, ao contrário da SVM. Os canais que foram selecionados para a LDA são apresentados nas Tabelas 32 e 33, os números de canais selecionados foram iguais, porém não foram os mesmos canais.

Tabela 30 – Quantidade de vezes que cada canal foi selecionado utilizando uma única SVM para o indivíduo B.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
1	FC5	1	AF3	1	PO7	0	FPz
1	FC3	1	AF4	1	PO3	0	AFz
1	FC1	1	F7	1	POz	0	AF8
1	FCz	1	F2	1	PO8	0	F5
1	FC2	1	F4	1	O1	0	F3
1	FC6	1	F8	1	Oz	0	F1
1	Cz	1	FT8	1	O2	0	Fz
1	C4	1	T8	1	Iz	0	F6
1	CP5	1	T9	0	FC4	0	FT7
1	CP1	1	TP7	0	C5	0	T7
1	CPz	1	TP8	0	C3	0	T10
1	CP2	1	P7	0	C1	0	P5
1	CP4	1	Pz	0	C2	0	P3
1	CP6	1	P2	0	C6	0	P1
1	FP2	1	P6	0	CP3	0	P4
1	AF7	1	P8	0	FP1	0	PO4

Tabela 31 – Número de caracteres classificados corretamente por uma única LDA utilizando diferentes combinações de canais.

64 Canais															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	12	15	25	28	31	42	51	53	55	60	58	60	68	71	73
B	28	39	46	53	61	66	71	74	78	80	80	83	84	85	86
Média	20	27	35,5	40,5	46	54	61	63,5	66,5	70	69	71,5	76	78	79,5
Fz,Cz,Pz, Oz, C3, C4, P3, P4, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	4	7	7	10	8	10	12	17	19	19	21	24	24	26	25
B	23	30	35	44	47	55	59	66	67	71	74	74	78	80	79
Média	13,5	18,5	21	27	27,5	32,5	35,5	41,5	43	45	47,5	49	51	53	52
PO7,Pz,CPz,P7, FC1,Cz, PO8 e FC5															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	4	6	10	12	14	14	16	22	23	27	32	34	40	43	44
B	23	25	39	41	47	54	59	64	68	72	73	76	78	78	78
Média	13,5	15,5	24,5	26,5	30,5	34	37,5	43	45,5	49,5	52,5	55	59	60,5	61
Seleção de Canais Iniciando com Pz, Oz, PO7 e PO8															
Indivíduo	Número de Sequências														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
A	9	18	25	24	28	33	37	45	47	50	49	55	58	62	65
B	28	38	49	57	65	69	75	75	77	80	81	81	81	81	84
Média	18,5	28	37	40,5	46,5	51	56	60	62	65	65	68	69,5	71,5	74,5

Tabela 32 – Quantidade de vezes que cada canal foi selecionado utilizando uma única LDA para o indivíduo A.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
1	FC5	1	CP4	1	F6	1	P6
1	FC3	1	CP6	1	F8	1	PO7
1	FC1	1	FP1	1	FT7	1	PO3
1	FCz	1	FPz	1	FT8	1	POz
1	FC2	1	FP2	1	T7	1	PO4
1	FC6	1	AF7	1	T8	1	PO8
1	C5	1	AF3	1	T9	1	O1
1	C3	1	AFz	1	T10	1	Oz
1	C1	1	AF4	1	TP7	1	O2
1	Cz	1	AF8	1	TP8	1	Iz
1	C4	1	F7	1	P7	0	FC4
1	C6	1	F3	1	P5	0	C2
1	CP5	1	F1	1	P3	0	CP1
1	CP3	1	Fz	1	P1	0	F5
1	CPz	1	F2	1	Pz	0	P4
1	CP2	1	F4	1	P2	0	P8

Tabela 33 – Quantidade de vezes que cada canal foi selecionado utilizando uma única LDA para o indivíduo B.

Quant.	Canais	Quant.	Canais	Quant.	Canais	Quant.	Canais
1	FC5	1	CPz	1	F8	1	P8
1	FC3	1	CP2	1	FT7	1	PO7
1	FC1	1	CP4	1	FT8	1	PO3
1	FC2	1	CP6	1	T7	1	POz
1	FC4	1	FPz	1	T8	1	PO4
1	FC6	1	AF7	1	T9	1	PO8
1	C5	1	AF3	1	T10	1	O1
1	C3	1	AFz	1	TP7	1	Oz
1	C1	1	AF8	1	TP8	1	O2
1	Cz	1	F7	1	P7	1	Iz
1	C2	1	F5	1	P5	0	FCz
1	C4	1	F3	1	P1	0	FP1
1	C6	1	F1	1	Pz	0	FP2
1	CP5	1	Fz	1	P2	0	AF4
1	CP3	1	F2	1	P4	0	F4
1	CP1	1	F6	1	P6	0	P3

Tabela 34 – Número de canais selecionados pela SVM e LDA.

Algoritmo	Indivíduos	
	A	B
SVM	64	40
LDA	58	58

## 4.7 Considerações Finais

Neste capítulo foi apresentada a metodologia utilizada para a detecção do P300 nos sinais de EEG e posterior identificação do caractere focado pelos usuários. A metodologia abordada envolve o pré-processamento dos sinais, segmentação, normalização, decimação, divisão da base de dados, treinamento, validação, teste dos modelos e os resultados. Foram discutidos fatores que influenciam nos resultados de cada um dos modelos e sugestões de modificações que podem melhorar o desempenho dos classificadores. Os resultados obtidos nos testes, e apresentados neste capítulo, mostraram que o desempenho foi melhor com o *ensemble* de SVMs com a seleção de canais. Em relação à seleção de canais, o *ensemble* de LDAs selecionou uma quantidade menor de canais, essa quantidade foi ainda menor com 5 LDAs. No entanto, o menor número de canais não resultou no melhor desempenho. O *ensemble* de LDAs tem como uma das vantagens o tempo de treinamento, e mesmo não atingindo o desempenho melhor que o *ensemble* de SVMs, ele é o mais indicado para ser utilizado em um soletrador.

---

## CONCLUSÕES

---

### 5.1 Contribuições e Limitações

Este trabalho realizou experimentos com a base de dados da *BCI Competition III Dataset II*. Esta base de dados contém dados gravados de 5 sessões de 2 indivíduos utilizando uma BCI baseada no P300 para controlar um soletrador virtual (*BCI speller*). Foram desenvolvidos vários modelos para identificar qual caractere cada um dos indivíduos desejava selecionar. Os modelos foram desenvolvidos a partir da variação de alguns parâmetros, tais como o algoritmo de aprendizado utilizado nos *ensembles*, SVM e LDA, assim como a quantidade de modelos utilizados. Variou-se também os canais que foram utilizados. Foram testados 8, 10 e 64 canais. Realizou-se também a seleção de canais em que incluía canais a partir de um conjunto inicial de canais pré-definidos.

Os resultados mostraram que, apesar da seleção de canais proposta selecionar mais canais que o método semelhante da literatura ([RAKOTOMAMONJY; GUIGUE, 2008](#)), houve uma redução significativa de canais e o desempenho com os números de sequências 1,2,3,4,5 e 10 foi melhor que alguns algoritmos propostos na literatura. Portanto, a hipótese do trabalho foi, em partes, confirmada. Os objetivos do trabalho foram atingidos. No entanto, a redução do tempo de treinamento dos classificadores está também associada ao algoritmo de aprendizado escolhido assim como à quantidade de exemplos. O tempo de treinamento das SVMs foi mais prejudicado pela quantidade de exemplos que o das LDAs, quanto mais exemplos eram utilizados, mais demorado era o treinamento. A dimensionalidade dos dados não apresentou efeitos perceptíveis no tempo de treinamento, mas tanto a quantidade como a dimensionalidade dos dados têm relação direta com o desempenho. Dependendo do algoritmo, o equilíbrio entre a quantidade e a dimensionalidade dos dados deve ser realizado com mais cautela.

A variação na quantidade de modelos do *ensemble*, com 1, 5 e 17 modelos, permitiu evidenciar algumas características de cada um dos algoritmos. A SVM se mostrou mais robusta

que a LDA quanto ao número de características utilizadas para treinamento. Porém, à medida que aumentava o número de exemplos de treinamento, a SVM apresentou diminuição de desempenho, em algumas configurações se tornou um classificador aleatório, enquanto que a LDA não foi tão prejudicada a este ponto.

A utilização dos conjuntos pré-definidos com 8 e 10 canais permitiu atingir desempenho médio de 84% com os dois indivíduos. No entanto, não atingiram resultados melhores que os obtidos com 64 canais e com a seleção de canais. O conjunto de canais iniciais para a seleção de canais pode ser alterado. Pode-se utilizar um dos dois conjuntos de canais pré-definidos que foram utilizados neste trabalho. É uma alternativa melhorar o desempenho e realizar o treinamento mais rápido, pois haverá menos canais para serem testados.

Uma das limitações dos modelos envolvendo a SVM é o tempo de treinamento e seleção de canais. Este processo pode ter o tempo reduzido de algumas formas. A seleção de canais pode usar como critério de parada um valor específico de avaliação. Durante os experimentos observou-se que as avaliações obtidas a partir da Equação 4.1 variaram entre 0 e 0,32, então a exploração de um valor adequado para ser utilizado como critério de parada agilizaria o processo de seleção de canais. Outra alternativa para diminuir o tempo de treinamento é diminuir a quantidade de exemplos de treinamento através da média dos sinais, como realizado em [Bhatnagar et al. \(2016\)](#), ou utilizando apenas sinais de alguns caracteres. Esta limitação está mais relacionada à SVM, mas poderia ser testada também com a LDA.

A decisão de como funcionaria o voto realizado nos *ensembles* também prejudicou um pouco o desempenho, pois foi identificado que ocorria, em uma pequena frequência, empate na identificação dos caracteres. Esse problema poderia ser diminuído, ou até menos eliminado, através de um voto ponderado.

## REFERÊNCIAS

---

---

- ABDI, H.; WILLIAMS, L. J. Principal component analysis. **Wiley interdisciplinary reviews: computational statistics**, Wiley Online Library, v. 2, n. 4, p. 433–459, 2010. Citado na página 40.
- ABDULHAY, E.; OWEIS, R.; MOHAMMAD, A.; AHMAD, L. Investigation of a wavelet-based neural network learning algorithm applied to P300 based brain-computer interface. **Biomedical Research**, Allied Academies, v. 29, p. 320–324, 2017. Citado nas páginas 39, 40, 51 e 57.
- AL-FAHOUM, A. S.; AL-FRAIHAT, A. A. Methods of eeg signal features extraction using linear analysis in frequency and time-frequency domains. **ISRN neuroscience**, Hindawi Publishing Corporation, v. 2014, 2014. Citado na página 38.
- AMIRI, S.; FAZEL-REZAI, R.; ASADPOUR, V. A review of hybrid brain-computer interface systems. **Advances in Human-Computer Interaction**, Hindawi Publishing Corp., v. 2013, p. 1, 2013. Citado na página 42.
- ARVANEH, M.; GUAN, C.; ANG, K. K.; QUEK, C. Optimizing the channel selection and classification accuracy in EEG-based BCI. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 58, n. 6, p. 1865–1873, 2011. Citado na página 41.
- AZLAN, W. A. W.; LOW, Y. F. Feature extraction of electroencephalogram (EEG) signal-A review. In: IEEE. **Biomedical Engineering and Sciences (IECBES), 2014 IEEE Conference on**. [S.l.], 2014. p. 801–806. Citado nas páginas 38, 40 e 61.
- BAŞAR-EROGLU, C.; BAŞAR, E.; DEMIRALP, T.; SCHÜRMMANN, M. P300-response: possible psychophysiological correlates in delta and theta frequency channels. a review. **International Journal of Psychophysiology**, Elsevier, v. 13, n. 2, p. 161–179, 1992. Citado na página 38.
- BHATNAGAR, V.; YEDE, N.; KERAM, R. S.; CHAURASIYA, R. A modified approach to ensemble of SVM for P300 based brain computer interface. In: IEEE. **Advances in Human Machine Interaction (HMI), 2016 International Conference on**. [S.l.], 2016. p. 1–5. Citado nas páginas 46, 50, 51, 64, 65, 66 e 76.
- BOSTANOV, V. Bci competition 2003-data sets Ib and IIb: feature extraction from event-related brain potentials with the continuous wavelet transform and the t-value scalogram. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 51, n. 6, p. 1057–1061, 2004. Citado na página 39.
- BURGES, C. J. A tutorial on support vector machines for pattern recognition. **Data mining and knowledge discovery**, Springer, v. 2, n. 2, p. 121–167, 1998. Citado na página 44.
- CASHERO, Z. Comparison of EEG preprocessing methods to improve the classification of P300 trials. **Master of Science Thesis, Colorado State University**, 2011. Citado nas páginas 40 e 57.

CORREIA, A. G. G. *et al.* Filtro notch para aplicações em EEGs e ECCs, com recurso a técnicas de F&H em cmos. 2010. Citado na página 38.

DEEDWANIYA, S.; GANDHI, T. K. An ensemble approach for brain computer interface applications. In: IEEE. **Humanitarian Technology Conference (R10-HTC), 2016 IEEE Region 10**. [S.l.], 2016. p. 1–4. Citado na página 51.

FARWELL, L. A.; DONCHIN, E. Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials. **Electroencephalography and Clinical Neurophysiology**, Elsevier, v. 70, n. 6, p. 510–523, 1988. Citado nas páginas 29 e 34.

FETZ, E. Operant conditioning of cortical unit activity. *Science*, v. 163, p. 955–958, 1969. Citado na página 29.

FLACH, P. **Machine learning: the art and science of algorithms that make sense of data**. [S.l.]: Cambridge University Press, 2012. Citado nas páginas 39 e 46.

FREEMAN, W.; QUIAN-QUIROGA, R. **Imaging brain function with EEG: advanced temporal and spatial analysis of electroencephalographic signals**. [S.l.]: Springer Science & Business Media, 2012. Citado na página 31.

FRIEDMAN, J.; HASTIE, T.; TIBSHIRANI, R. **The elements of statistical learning**. [S.l.]: Springer series in statistics New York, NY, USA:, 2001. v. 1. Citado na página 45.

GARRETT, D.; PETERSON, D. A.; ANDERSON, C. W.; THAUT, M. H. Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, Citeseer, v. 11, n. 2, p. 141–144, 2003. Citado na página 45.

HE, B.; GAO, S.; YUAN, H.; WOLPAW, J. R. Brain–computer interfaces. In: **Neural Engineering**. [S.l.]: Springer, 2013. p. 87–151. Citado nas páginas 30, 32 e 33.

JAIN, A. K.; DUIN, R. P.; MAO, J. Statistical pattern recognition: A review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Ieee, v. 22, n. 1, p. 4–37, 2000. Citado na página 40.

JUNG, T.-P.; MAKEIG, S.; HUMPHRIES, C.; LEE, T.-W.; MCKEOWN, M. J.; IRAGUI, V.; SEJNOWSKI, T. J. Removing electroencephalographic artifacts by blind source separation. **Psychophysiology**, Cambridge University Press, v. 37, n. 2, p. 163–178, 2000. Citado na página 38.

KAPER, M.; MEINICKE, P.; GROSSEKATHOEFER, U.; LINGNER, T.; RITTER, H. BCI competition 2003-data set IIb: support vector machines for the P300 speller paradigm. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 51, n. 6, p. 1073–1076, 2004. Citado nas páginas 15, 44 e 51.

KEE, C.-Y.; PONNAMBALAM, S.; LOO, C.-K. Multi-objective genetic algorithm as channel selection method for P300 and motor imagery data set. **Neurocomputing**, Elsevier, v. 161, p. 120–131, 2015. Citado na página 41.

KRUSIENSKI, D. J.; SELLERS, E. W.; MCFARLAND, D. J.; VAUGHAN, T. M.; WOLPAW, J. R. Toward enhanced P300 speller performance. **Journal of Neuroscience Methods**, Elsevier, v. 167, n. 1, p. 15–21, 2008. Citado nas páginas 40 e 57.



- KUNDU, S.; ARI, S. P300 detection with brain–computer interface application using PCA and ensemble of weighted SVMs. **IETE Journal of Research**, Taylor & Francis, v. 64, n. 3, p. 406–414, 2018. Citado nas páginas 46, 51, 53, 54 e 64.
- KUSUMANDARI, D. E.; HIDAYAT, T.; TURNIP, A. A Comparison of Real-Time Extraction between Chebyshev and Butterworth Method for SSVEP Brain Signals. **Journal of Telecommunication, Electronic and Computer Engineering (JTEC)**, v. 10, n. 2-3, p. 135–139, 2018. Citado na página 39.
- LAL, T. N.; SCHRODER, M.; HINTERBERGER, T.; WESTON, J.; BOGDAN, M.; BIRBAUMER, N.; SCHOLKOPF, B. Support vector channel selection in BCI. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 51, n. 6, p. 1003–1010, 2004. Citado nas páginas 40 e 41.
- LEZA, C.; PUTHUSSERYPADY, S. Detection of user independent single trial ERPs in brain computer interfaces: An adaptive spatial filtering approach. In: IEEE. **Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE**. [S.l.], 2017. p. 2243–2246. Citado na página 25.
- LI, Q.; SHI, K.; MA, S.; GAO, N. Improving classification accuracy of SVM ensemble using random training set for BCI P300-speller. In: IEEE. **Mechatronics and Automation (ICMA), 2016 IEEE International Conference on**. [S.l.], 2016. p. 2611–2616. Citado na página 39.
- LORENA, A. C.; CARVALHO, A. C. de. Uma introdução às support vector machines. **Revista de Informática Teórica e Aplicada**, v. 14, n. 2, p. 43–67, 2007. Citado na página 44.
- LOTTE, F.; CONGEDO, M.; LÉCUYER, A.; LAMARCHE, F.; ARNALDI, B. A review of classification algorithms for EEG-based brain–computer interfaces. **Journal of Neural Engineering**, IOP Publishing, v. 4, n. 2, p. R1, 2007. Citado na página 42.
- LUCK, S. J. **An introduction to the event-related potential technique**. [S.l.]: MIT press, 2014. Citado na página 32.
- MARSHALL, D.; COYLE, D.; WILSON, S.; CALLAGHAN, M. Games, gameplay, and BCI: the state of the art. **IEEE Transactions on Computational Intelligence and AI in Games**, IEEE, v. 5, n. 2, p. 82–99, 2013. Citado na página 25.
- MARTÍNEZ-CAGIGAL, V.; SANTAMARÍA-VÁZQUEZ, E.; HORNERO, R. A novel hybrid swarm algorithm for P300-based BCI channel selection. In: SPRINGER. **World Congress on Medical Physics and Biomedical Engineering 2018**. [S.l.], 2019. p. 41–45. Citado na página 41.
- MILLÁN, J. d. R.; RUPP, R.; MÜLLER-PUTZ, G. R.; MURRAY-SMITH, R.; GIUGLIEMMA, C.; TANGERMANN, M.; VIDAURRE, C.; CINCOTTI, F.; KÜBLER, A.; LEEB, R.; NEUPER, C.; MUELLER, K. R.; MATTIA, D. Combining brain–computer interfaces and assistive technologies: state-of-the-art and challenges. Frontiers Research Foundation, 2010. Citado na página 25.
- MOAZZAMI, M.-M. **EEG signal Processing in Brain-Computer Interface**. Tese (Doutorado) — Michigan State University, 2012. Citado nas páginas 32 e 33.
- NIJBOER, F.; SELLERS, E.; MELLINGER, J.; JORDAN, M. A.; MATUZ, T.; FURDEA, A.; HALDER, S.; MOCHTY, U.; KRUSIENSKI, D.; VAUGHAN, T. *et al.* A P300-based brain–computer interface for people with amyotrophic lateral sclerosis. **Clinical Neurophysiology**, Elsevier, v. 119, n. 8, p. 1909–1916, 2008. Citado nas páginas 25 e 26.

NIJHOLT, A.; TAN, D. S. **Brain-Computer Interfaces: Applying our Minds to Human-Computer Interaction**. [S.l.]: Springer, 2010. Citado nas páginas 37 e 38.

OBERMAIER, B.; MULLER, G.; PFURTSCHELLER, G. "Virtual keyboard" controlled by spontaneous EEG activity. **IEEE Transactions on Neural Systems and Rehabilitation Engineering**, IEEE, v. 11, n. 4, p. 422–426, 2003. Citado na página 30.

PERTENCE, J. A. **Amplificadores Operacionais e Filtros Ativos-8**. [S.l.]: Bookman Editora, 2015. Citado na página 39.

PFURTSCHELLER, G.; NEUPER, C. Motor imagery and direct brain-computer communication. **Proceedings of the IEEE**, v. 89, n. 7, p. 1123–1134, 2001. Citado na página 30.

RAKOTOMAMONJY, A.; GUIGUE, V. BCI competition III: dataset II-ensemble of SVMs for BCI P300 speller. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 55, n. 3, p. 1147–1154, 2008. Citado nas páginas 19, 39, 40, 41, 42, 46, 50, 51, 52, 53, 54, 58, 60, 61, 64, 65, 68 e 75.

RAMADAN, R. A.; REFAT, S.; ELSHAHED, M. A.; ALI, R. A. Basics of brain-computer interface. In: **Brain-Computer Interfaces**. [S.l.]: Springer, 2015. p. 31–50. Citado nas páginas 15, 25, 26, 31, 33 e 34.

RAMEZANI, M.; MARBLE, K.; TRANG, H.; JOHNSRUDE, I. S.; ABOLMAESUMI, P. Joint sparse representation of brain activity patterns in multi-task fmri data. **IEEE transactions on medical imaging**, IEEE, v. 34, n. 1, p. 2–12, 2015. Citado na página 52.

SCHAPIRE, R. E. The boosting approach to machine learning: An overview. In: **Nonlinear Estimation and Classification**. [S.l.]: Springer, 2003. p. 149–171. Citado na página 46.

SELLERS, E. W.; VAUGHAN, T. M.; WOLPAW, J. R. A brain-computer interface for long-term independent home use. **Amyotrophic Lateral Sclerosis**, Taylor & Francis, v. 11, n. 5, p. 449–455, 2010. Citado na página 40.

SPEIER, W.; DESHPANDE, A.; POURATIAN, N. A method for optimizing EEG electrode number and configuration for signal acquisition in p300 speller systems. **Clinical Neurophysiology**, Elsevier, v. 126, n. 6, p. 1171–1177, 2015. Citado na página 26.

SUR, S.; SINHA, V. Event-related potential: An overview. **Industrial Psychiatry Journal**, Medknow Publications, v. 18, n. 1, p. 70, 2009. Citado nas páginas 33 e 34.

SZAFIR, D. Non-invasive BCI through eeg. **Undergraduate Honors Thesis in Computer Science, Boston College**, 2010. Citado na página 32.

THARWAT, A.; GABER, T.; IBRAHIM, A.; HASSANIEN, A. E. Linear discriminant analysis: A detailed tutorial. **AI Communications**, IOS Press, v. 30, n. 2, p. 169–190, 2017. Citado na página 45.

VALLABHANENI, A.; WANG, T.; HE, B. Brain—computer interface. In: **Neural engineering**. [S.l.]: Springer, 2005. p. 85–121. Citado nas páginas 29, 33, 34 e 38.

VAPNIK, V. **The nature of statistical learning theory**. [S.l.]: Springer science & business media, 2013. Citado nas páginas 44 e 66.

VIDAL, J. J. Toward direct brain-computer communication. **Annual review of Biophysics and Bioengineering**, Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, CA 94303-0139, USA, v. 2, n. 1, p. 157–180, 1973. Citado na página 29.

\_\_\_\_\_. Real-time detection of brain events in EEG. **Proceedings of the IEEE**, IEEE, v. 65, n. 5, p. 633–641, 1977. Citado na página 29.

WANG, D.; PENG, J. Wavelet analysis and application in signal processing. **Beijing: Publishing House of Electronic Industry**, 2006. Citado nas páginas 40 e 57.

WOLPAW, J. R.; BIRBAUMER, N.; MCFARLAND, D. J.; PFURTSCHELLER, G.; VAUGHAN, T. M. Brain-computer interfaces for communication and control. **Clinical Neurophysiology**, Elsevier, v. 113, n. 6, p. 767–791, 2002. Citado nas páginas 15, 30, 31, 32, 34 e 57.

XU, N.; GAO, X.; HONG, B.; MIAO, X.; GAO, S.; YANG, F. Bci competition 2003-data set IIb: enhancing P300 wave detection using ICA-based subspace projections for BCI applications. **IEEE Transactions on Biomedical Engineering**, IEEE, v. 51, n. 6, p. 1067–1072, 2004. Citado na página 40.

