



EDSON GEOVANI DOS SANTOS

**ANÁLISE DE TREINAMENTO E CLASSIFICAÇÃO
DE SINAL EEG DE EXECUÇÃO DE MEMBROS INFERIORES
COLETADOS POR BCI VISANDO DESEMPENHO
E REDUÇÃO DE CANAIS**

**MESTRADO EM
TECNOLOGIA EM SAÚDE
PUCPR**

**CURITIBA
2023**

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE**

**ANÁLISE DE TREINAMENTO E CLASSIFICAÇÃO DE SINAL EEG DE
EXECUÇÃO DE MEMBROS INFERIORES COLETADOS POR BCI VISANDO
DESEMPENHO E REDUÇÃO DE CANAIS**

PROJETO DE DEFESA DE MESTRADO

EDSON GEOVANI DOS SANTOS

**ORIENTADOR: Prof. Dr. GUILHERME NUNES NOGUEIRA NETO
COORIENTADOR: Prof. Dr. PERCY NOHAMA**

**CURITIBA
2023**

EDSON GEOVANI DOS SANTOS

**ANÁLISE DE TREINAMENTO E CLASSIFICAÇÃO DE SINAL EEG DE
EXECUÇÃO DE MEMBROS INFERIORES COLETADOS POR BCI VISANDO
DESEMPENHO E REDUÇÃO DE CANAIS**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Tecnologia em Saúde. Curso de pós-graduação em tecnologia e saúde (PPGTS) – Pontifícia Universidade Católica do Paraná (PUCPR)
Orientador: Prof. Dr. Guilherme Nunes Nogueira Neto
Coorientador: Prof. Dr. Percy Nohama

**CURITIBA
2023**

Dados da Catalogação na Publicação
Pontifícia Universidade Católica do Paraná
Sistema Integrado de Bibliotecas – SIBI/PUCPR
Biblioteca Central
Edilene de Oliveira dos Santos CRB-9 /1636

S237a Santos, Edson Geovani dos
2023 Análise de treinamento e classificação de sinal EEG de execução de membros inferiores coletados por BCI visando desempenho e redução de canais / Edson Geovani dos Santos ; orientadora: Guilherme Nunes Nogueira Neto ; coorientador: Percy Nohama. -- 2023
69 f. ; il. : 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2023
Bibliografia: f. 64-69

1.Tecnologia médica. 2. Eletroencefalograma. 3. Interfaces cérebro - Computador. 4. Aprendizado de máquina. 5. Extremidade inferior. I. Nogueira Neto, Guilherme Nunes. II. Nohama, Percy. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Tecnologia em Saúde. IV. Título.

CDD. 20. ed. – 610.28



TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 317

A Dissertação de Mestrado intitulada: **“ANÁLISE DE TREINAMENTO E CLASSIFICAÇÃO DE SINAL EEG DE EXECUÇÃO DE MEMBROS INFERIORES COLETADOS POR BCI VISANDO DESEMPENHO E REDUÇÃO DE CANAIS”**, defendida em sessão pública pelo(a) candidato(a) **Edson Geovani dos Santos** em 28 de fevereiro de 2023, foi julgada para a obtenção do grau de Mestre em Tecnologia em Saúde, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Tecnologia em Saúde.

BANCA EXAMINADORA:

Prof. Dr. Guilherme Nunes Nogueira Neto – (PUCPR)

Prof. Dr. Alceu de Souza Britto Junior – (PUCPR)

Prof. Dr. André Eugênio Lazzaretti - (UTFPR)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 15 de junho de 2023.

Prof. Dr. Percy Nohama
**Coordenador do Programa de
Pós-Graduação em Tecnologia em Saúde
PUCPR**

Dedico este trabalho a todos os que me
ajudaram ao longo desta jornada.

AGRADECIMENTOS

Agradeço primeiramente a Deus, o Qual me proporcionou a oportunidade de estudar e chegar até esse ponto. Agradeço também a instituição PUC, a qual tive oportunidade de estudar e ser reconhecido por meio do prêmio Marcelino Champagnat, e com este ingressar no programa de pós-graduação.

Agradeço a todos os professores responsáveis pela minha aceitação no programa, especialmente aos Professor Guilherme Nunes Nogueira Neto e ao Professor Percy Nohama, os quais estiveram comigo me orientando no trajeto desta dissertação. Um agradecimento especial à Professora Dra. Paula Hembecker, a qual me motivou e ajudou muito na reta final.

Um agradecimento mais que especial a Andressa Luisa Sensolo, que esteve ao meu lado todo esse tempo, prestando suporte motivacional e emocional, a qual sem ela, possivelmente, eu não estaria aqui para prestigiar você leitor com este trabalho.

Por fim, agradeço a minha família, que sempre me incentivou, facilitou e encorajou que eu me dedicasse aos meus estudos, provendo o necessário para que atingisse meus objetivos.

RESUMO

Introdução: Autores afirmam que a preservação de capacidades funcionais motoras é o que tem levado à determinação de um indivíduo ser saudável ou não. Diversas doenças, lesões nervosas, neuromusculares ou osteoarticulares, e até mesmo má formação, podem interferir nas capacidades funcionais de um indivíduo, afetando diretamente sua qualidade de vida. Para preservação da autonomia dos indivíduos acometidos por essas deficiências o uso de aprendizado de máquina no treinamento de sinais EEG coletados por BCIs têm ganhado atenção, pois permite o acionamento de dispositivos assistivos utilizando sinais cerebrais que não dependem do acionamento muscular, como órteses e próteses que auxiliam na recuperação de movimentos. **Objetivos:** Investigar combinações de treinamentos com algoritmos de aprendizado de máquina para a classificação de sinais de EEG em tarefas de membros inferiores, visando aquelas com o menor número de canais e melhor desempenho. **Método:** Foi realizada, inicialmente, uma revisão da literatura para definir as diretrizes para os testes conduzidos neste estudo. Foram realizados 1407 treinamentos com sinais de EEG de uma base pública, empregando os algoritmos SVM, LDA e kNN. Variou-se a quantidade de canais, além das quantidades de exemplos de treinamento. Os algoritmos foram codificados no MATLAB, onde os resultados de exatidão para cada combinação foram calculados e analisados. **Resultados:** A partir da prospecção da literatura, determinou-se SVM, LDA e kNN e as configurações de canais com melhor performance e, a partir dos treinamentos, a combinação de ML x quantidade de canais que apresentou melhor desempenho foi SVM com 1 Canal (Cz). O impacto na predição de novas tarefas apresentadas ao classificador atingiu até 88,30% de exatidão para reconhecimento de 60 novas tarefas. **Conclusões:** Assim como visto na revisão da literatura, o classificador que obteve o melhor desempenho acerca de todos os testes realizados foi o SVM, porém a configuração do canal Cz diferiu da literatura, tendo em vista que neste estudo se apresentou com melhor desempenho do que quando comparado ao Fp1, que obteve maior exatidão média na literatura. A partir disso, é possível afirmar que é possível prever novos movimentos mesmo com apenas 1 canal, mantendo níveis de desempenho praticamente equivalentes aos de quando aplicado a mais canais. Além disso, foi visto que quanto mais exemplos de tarefas realizadas durante o treinamento melhor será a exatidão da máquina ao reconhecer movimentos de membros inferiores. **Palavras-chave:** EEG. Interface Cérebro-Máquina. Membros Inferiores. Aprendizado de Máquina. Redução de Canais.

ABSTRACT

Introduction: Authors state that the preservation of motor functional capacities is what has led to the determination of an individual to be healthy or not. Several diseases, nerve, neuromuscular or osteoarticular injuries, and even fetus malformation, can interfere with an individual's functional capabilities, directly affecting their quality of life. To preserve the autonomy of individuals affected by these disabilities, the use of machine learning in training EEG signals collected by BCIs has gained attention, as it allows the activation of assistive devices using brain signals that do not depend on muscular activation, such as orthoses and prostheses that aid in movement recovery.

Objectives: To investigate combination training with machine learning algorithms for classifying EEG signals in lower limb tasks, targeting those with the fewest channels and best performance.

Method: Initially, a literature review was carried out to define the guidelines for the tests conducted in this study. 1407 trainings were carried out with EEG signals from a public database, using the SVM, LDA and kNN algorithms. The number of channels was varied, in addition to the number of training examples. The algorithms were coded in MATLAB, where the accuracy results for each combination were calculated and analyzed.

Results: From the literature survey, SVM, LDA, and kNN and the channel configurations with the best performance were determined, and from the trainings, the combination of ML x channel quantity that showed the best performance was SVM with 1 Channel (Cz). The impact on the prediction of new tasks presented to the classifier reached up to 88.30% accuracy for recognition of 60 new tasks.

Conclusions: As seen in the literature review, the classifier that obtained the best performance regarding all the tests performed was the SVM, but the Cz channel configuration differed from the literature, given that in this study it presented with better performance than when compared to Fp1, which obtained a higher average accuracy in the literature. Based on this, it is possible to affirm that it is possible to predict new movements even with only 1 channel, maintaining levels of performance almost equivalent to those when applied to more channels. Furthermore, it was seen that the more examples of tasks performed during training, the better the accuracy of the machine in recognizing lower limb movements.

Keywords: EEG. Brain-Computer Interface. Lower Limbs. Machine Learning. Reduction of Channel Selection.

LISTA DE FIGURAS

Figura 1 - Taxonomia abstrata do desenvolvimento de BCIs	21
Figura 2 - Regiões do cérebro	22
Figura 3 - Sistemas de Posicionamento Internacional	23
Figura 4 - Representação gráfica de uma validação cruzada 10-fold	27
Figura 5 - Métodos de Machine Learning disponíveis no MATLAB	30
Figura 6 - Fluxograma do PRISMA descrevendo as etapas da revisão e artigos selecionados	31
Figura 7 - Mapa de calor dos eletrodos mais utilizados encontrados na literatura ...	35
Figura 8 - Diagrama resumindo a metodologia utilizada	38
Figura 9 - Representação gráfica do sinal EEG	39
Figura 10 - Representação gráfica de sinal EEG enfatizando os valores das amostras coletadas.....	40
Figura 11 - Configuração padrão do SVM	42
Figura 12 - Configuração padrão do LDA	42
Figura 13 - Configuração padrão do kNN	43

LISTA DE TABELAS

Tabela 1 - Matriz de Confusão	28
Tabela 2 - Quantidade e seleção de canais escolhidos	34
Tabela 3 - Resultados médios obtidos a partir dos treinamentos realizados para 30 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais e 15 sujeitos individuais	47
Tabela 4 - Resultados médios obtidos a partir dos treinamentos realizados para 90 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais e 15 sujeitos individuais	47
Tabela 5 - Resultados médios obtidos a partir dos treinamentos realizados para 90 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais, 15 sujeitos coletivos e 105 sujeitos coletivos	48
Tabela 6 - Média de exatidão para cada classificador; Acc = exatidão; DP = Desvio Padrão.....	48
Tabela 7 - Classificação dos resultados médios por quantidade de canais para todos os classificadores. Mín. Acc = Menor valor de exatidão obtido; Máx. Acc = Maior valor de exatidão obtido	49
Tabela 8 - Classificação dos resultados médios por quantidade de canais para o classificador SVM. Mín. Acc = Menor valor de exatidão obtido; Máx. Acc = Maior valor de exatidão obtido	50
Tabela 9 - Valores máximos de exatidão obtidos pelo treinamento de um canal (Cz) para cada indivíduo para o algoritmo SVM	51
Tabela 10 - Treinamento realizado com o sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 46 tarefas	53
Tabela 11 - Treinamento realizado com o sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 90 tarefas	53
Tabela 12 – Teste da base de dados do sujeito S015 executando 90 tarefas realizado com o treinamento realizado para sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 46 tarefas	54

LISTA DE ACRÔNIMOS

ACC – Exatidão

BCI – Brain-Computer Interface

DP – Desvio Padrão

EEG – Electroencephalogram

FN – Falso Negativo

FNR – False Negative Rate

FP – Falso Positivo

FPR – False Positive Rate

kNN – k-Nearest Neighbor

LDA – Linear Discriminant Analysis

Máx. Acc. – Maior valor de exatidão obtido

Mín. Acc. – Menor valor de exatidão obtido

ML – Machine Learning

RLDA – Regularized Linear Discriminant Analysis

SVM – Support Vector Machine

TNR – True Negative Rate

TPR – True Positive Rate

TRTR – Treino e teste da base de dados real

TRTS – Treino na base real e teste na base sintética

TSTR – Treino na base sintética e teste na base real

TSTS – Treino e teste na base sintética

VN – Verdadeiro Negativo

VP – Verdadeiro Positivo

μV – *Microvolt*

SUMÁRIO

1. INTRODUÇÃO	13
1.1. APRESENTAÇÃO	14
1.1.1. Problemática	16
1.1.2. Justificativa	17
1.2. OBJETIVOS	18
1.2.1. Objetivo geral	18
1.2.2. Objetivos específicos	18
2. FUNDAMENTAÇÃO TEÓRICA	20
2.1. INTERFACE CÉREBRO-MÁQUINA E ELETROENCEFALOGRAMA	20
2.2. REGIÕES DO CÉREBRO	21
2.3. SISTEMA DE POSICIONAMENTO INTERNACIONAL	22
2.4. APRENDIZADO DE MÁQUINA	23
2.5. CLASSIFICADORES	25
2.5.1. Support Vector Machine	25
2.5.2. Linear Discriminant Analysis	26
2.5.3. K-Nearest Neighbor	26
2.5.4. Validação cruzada <i>k-fold</i>	27
2.5.5. Métricas de avaliação	28
2.6. MATLAB	29
3. ESTADO DA ARTE	31
4. MATERIAIS E MÉTODOS	38
4.1. BASE DE DADOS	38
4.2. HARDWARE E SOFTWARE	41
4.3. TREINAMENTOS	43
5. RESULTADOS	47
5.1. TAREFA 1	48
5.2. TAREFA 2	50
5.3. TAREFA 3	52
5.4. TAREFA 4	52
5.5. TAREFA 5	53
6. DISCUSSÃO	55
7. CONCLUSÕES	61
8. CONTRIBUIÇÕES	62
9. TRABALHOS FUTUROS	63

REFERÊNCIAS BIBLIOGRÁFICAS.....	64
--	-----------

1. INTRODUÇÃO

Existem doenças que podem originar fraquezas musculares nos membros inferiores. De acordo com a ONU (2021), na União Europeia, incidentes de derrames totalizam aproximadamente um milhão por ano, e as sequelas podem variar de diversos comprometimentos ou déficits neurológicos até hemiplegia, afetando as funcionalidades dos membros inferiores.

Uma forma de tentar minimizar os efeitos das disfunções motoras é empregando tecnologias assistivas em mobilidade, que são serviços, recursos ou práticas que fomentam a máxima participação de um indivíduo com limitações específicas (BERSCH, 2009).

Existem diversos exemplos de estudos, como Dos Santos; Paz (2018), Pratt, Collins, Krupp *et al.* (2004) e Xu *et al.* (2014), que utilizam mecanismos manuais ou programados para ambiente de laboratório na realização de movimentos específicos e implicam o início do sistema por meio de acionadores tanto por parte de técnicos, quanto o próprio paciente. Assim, percebe-se a necessidade de aprimoramentos e aplicação de um sistema automatizado/inteligente que envolva tecnologias assistivas para a mobilidade dos usuários. Controlar dispositivos usando a mente tem fascinado a humanidade e no contexto da proposta, uma forma de realizar isso é por meio do acionamento de órteses a partir da atividade cerebral.

Um sistema de Interface Cérebro-Máquina (BCI – Brain-Computer Interfaces) consiste em diversos processamentos de sinais e sensores que permitem a análise da atividade cerebral. Em síntese, a atividade cerebral é captada por meio de eletrodos aplicados no escalpo (mais comumente aplicado no formato não invasivo) e os sinais coletados (conhecidos como eletroencefalograma - EEG) podem ser processados e estudados com a utilização de outros recursos (BIRBAUMER; COHEN, 2007).

Dentre as possíveis aplicações para o emprego das BCIs, destaca-se a oportunidade para controlar dispositivos a partir da atividade cerebral, ou seja, eles possuem a capacidade de ler as atividades cerebrais e traduzir as mesmas em comandos de controle (BLANKERTZ *et al.*, 2006).

O que concerne às disfunções motoras, uma característica de interesse é que as BCIs disponibilizam ao usuário canais de controle e comunicação que não dependem dos resultados obtidos pelos músculos e nervos periféricos. Isso torna o desenvolvimento dessa tecnologia mais atrativa, pois espera-se obter aumento das opções de comunicação para os indivíduos com severas deficiências motoras, as quais impedem a utilização de tecnologias auxiliares convencionais (WOLPAW *et al.*, 2006). Sendo assim, esses sistemas possuem grande potencial de apresentar benefícios para as pessoas com deficiência física, sendo os principais, a melhoria na qualidade de vida e maior autonomia (KUBLER *et al.*, 2006).

Quando utilizado um BCI para coletar sinais de EEG, são empregados alguns classificadores, que são uma forma de identificar algumas características do sinal e assim realizar estudos e/ou aplicações. Os classificadores são um tipo de aprendizado de máquina (ML), sendo este um dos ramos da computação que busca desenvolver e treinar máquinas capazes de simular a inteligência humana a partir de seu próprio aprendizado, muitas vezes sendo baseado em reconhecimento de padrões aplicado em diversas áreas desde finanças até aplicações em medicina (LEE, EL NAQA, 2015).

Em conclusão, as tecnologias assistivas em mobilidade são serviços, recursos ou práticas que fomentam a máxima participação de indivíduos com limitações específicas. A interface Cérebro-Máquina (BCI) é uma tecnologia que tem o potencial de ajudar pessoas idosas ou com disfunções motoras, pois permite o controle de dispositivos com o uso da atividade cerebral. Estudos mostram a necessidade de aprimoramentos e aplicação de sistemas automatizados/inteligentes para auxiliar na mobilidade dos usuários. É importante continuar investigando e desenvolvendo essas tecnologias para melhorar a qualidade de vida e a independência dessas pessoas.

1.1. APRESENTAÇÃO

A estrutura desta dissertação foi separada em capítulos, sendo o primeiro, um capítulo introdutório separado em apresentação, que consiste em uma sessão para apresentar a estrutura, problemática e justificativa do trabalho, e objetivos, apresentando os objetivos gerais e específicos.

O segundo capítulo consiste na fundamentação teórica, a qual consiste em uma exposição dos principais conceitos utilizados durante a dissertação. Dentro deste

capítulo são explanados os conceitos de Inferace Cérebro-Máquina (*Brain-Computer Interface*), das regiões do cérebro, do sistema de posicionamento internacional de eletrodos, de extratores de características de sinal, de aprendizado de máquina, de classificadores e do software MATLAB.

Além disso, a sessão de classificadores é subdividida em *Support Vector Machine*, *Linear Discriminant Analysis*, *k-Nearest Neighbor*, Validação cruzada *k-fold* e métricas de avaliação.

O terceiro capítulo consiste no estado da arte, o qual expõem os principais documentos encontrados na literatura relacionados aos assuntos tratados nessa dissertação. Esses documentos foram localizados através da realização de uma revisão sistemática, e todo o processo de realização desta também é descrito nessa sessão.

O quarto capítulo trata dos materiais e métodos utilizados para o desenvolvimento deste trabalho. Primeiramente é apresentada a base de dados utilizada para a como os seus autores a coletaram. Em sequência, apresenta-se a configuração de hardware e software utilizada detalhadamente.

Por fim, é apresentada uma sessão de treinamentos que expõem minuciosamente como foram realizados os preparativos para execução dos treinamentos e testes dos classificadores utilizados. Além disso, explica-se, em separado por tarefa, quais eram as ações que seriam realizadas durante o projeto.

O quinto capítulo consiste na apresentação dos resultados, o qual está separado para apresentar os resultados de cada tarefa descrita anteriormente. O sexto capítulo consiste nas discussões dos resultados apresentados com os demais autores encontrados na literatura.

Em sequência estão os capítulos 7, sendo esse a exposição das conclusões obtidas a partir dos resultados e discussões, capítulo 8, apresentando as contribuições que o trabalho trouxe para a literatura, e o capítulo 9, o qual consiste em apresentar os possíveis trabalhos futuros.

1.1.1. Problemática

De acordo com a Política Nacional de Saúde da Pessoa com Deficiência (2017), pessoas com incapacidades motoras podem ser acometidas por uma variedade de condições neurossensoriais que as afetam em termos de mobilidade ou coordenação motora geral devido a lesões nervosas, neuromusculares, osteoarticulares ou má formação.

Conforme o Professor Doutor Luzimar Teixeira (2010), a incapacidade motora pode ser entendida por alterações corporais que provocam dificuldades na movimentação das pessoas e às impedem de participar na sociedade de forma independente. Pode ser classificada em distúrbios ortopédicos, que originam em músculos, articulações ou ossos, ou neurológicos, que acontecem no sistema nervoso. Além disso, o professor explica que podem ser definidas em congênitas ou adquiridas, agudas ou crônicas, permanentes ou temporárias e progressivas ou não progressivas. Existem diversas causas que podem originar as incapacidades motoras, como problemas genéticos, complicações na gestação, doenças e acidentes.

Outra forma de desenvolver a incapacidade muscular é através da sarcopenia, que consiste na redução da capacidade funcional de pessoas idosas, que causa naturalmente a diminuição da força muscular, da flexibilidade e da resistência aeróbia, com isso dificulta a realização de ações cotidianas como caminhar, subir escadas e levantar-se de cadeiras (DIAS; GURJÃO; MARUCCI, 2006).

De acordo com Bonardi, Souza e Moraes (2007), a definição de uma pessoa saudável não pode mais ser considerada a partir de quantas doenças um indivíduo possui, mas do grau de preservação de suas capacidades funcionais, incluindo os estados fisiológico e psicológico.

A degradação destes aspectos em idosos resulta na dificuldade de manter a homeostasia corporal, ou seja, dificuldade em manter sua estabilidade do meio interno corporal, durante o avanço de patologias crônicas degenerativas, que podem gerar tais ineficiências musculares que diminuem a qualidade de vida das pessoas acometidas.

1.1.2. Justificativa

Levando em consideração que diversas limitações são causadas pela ineficiência muscular, evidenciou-se que é válido e necessário buscar meios que proporcionem às pessoas com fraquezas em membros inferiores e doenças crônicas uma melhora na qualidade de vida, tendo em vista o limitado acervo de estudos técnicos que a englobam. Um dos métodos é a integração de sistemas de BCI com técnicas de inteligência artificial que possibilitam uma perspectiva para o acionamento de dispositivos de tecnologias assistivas.

Assim, a atividade cerebral poderia acionar de forma automática e inteligente os atuadores que movimentam as articulações de uma órtese, livrando o indivíduo da necessidade de pressionar um botão ou emitir um comando de voz. Portanto, tendo em vista o exposto, surge o interesse em estudar condições que facilitem a aprimorem o acesso dessas tecnologias assistivas por meio da redução de canais para leitura de sinais EEG de movimentos em membros inferiores, tendo em vista que menos canais trariam uma grande redução de custo já no momento de aquisição do equipamento.

Além disso, como o objetivo dessas tecnologias assistivas é classificar o movimento quando a ação é iniciada no cérebro do indivíduo que a utiliza, sendo assim, faz-se necessário um sistema de processamento e de classificação que gere um retorno rápido e preciso, justificando assim a necessidade de estudar quais os métodos já existentes e mais comuns utilizados pelo mundo acadêmico e verificar se a eficiência deles pode corresponder às expectativas esperadas além de verificar se os mesmos apresentam a velocidade de resposta desejada.

Tendo em vista que o custo de tecnologias assistivas não é acessível para a maior parte da população, a redução de canais entraria como uma forma de baratear os custos de desenvolvimento futuro dessa tecnologia, da forma que quando mais canais, mais caros são encontrados os equipamentos BCIs. Sendo assim, a redução de canais é verificada para aumentar a velocidade de processamento e baratear custos.

Portanto, faz-se necessário também a realização de uma pesquisa de estado da arte referente à aplicação de BCIs para a coleta de sinais EEG em adultos e classificação, por meio de algoritmos de inteligência de máquina, de movimentos de

membros inferiores para direcionar quais os classificadores mais utilizados na literatura, com a maior exatidão já obtida. Além disso, verificar os estudos que já foram realizados em relação à redução de canais e entender qual é o classificador mais indicado para uma menor quantidade de informações.

Outro fator importante é a determinação de quais escolhas de canais tiveram os maiores valores de exatidão para cada quantidade de canais selecionados, podendo assim ter um *baseline* pré-determinado direcionando o andamento da pesquisa.

Por fim, pode-se buscar uma base de dados renomada de sinais EEG de movimento de membros inferiores para então realizar o treinamento desses sinais com os classificadores mais utilizados e com os melhores resultados encontrados na literatura e verificar se eles realmente apresentam uma boa exatidão, além de ter boa exatidão quando os canais utilizados para a coleta são reduzidos.

1.2. OBJETIVOS

Levando em consideração a problemática apresentada, foram propostos os objetivos gerais e específicos para a realização desta dissertação, descritos a seguir.

1.2.1. Objetivo geral

Investigar combinações de treinamentos com algoritmos de aprendizado de máquina para a classificação de sinais de EEG em tarefas de membros inferiores, visando aquelas com o menor número de canais e melhor desempenho.

1.2.2. Objetivos específicos

1. Identificar algoritmos de aprendizado de máquina com os melhores índices de desempenho, bem como os canais utilizados para classificação de sinais EEG, prospectando na literatura sobre BCIs baseadas em sinais EEGs para movimentos de membros inferiores;
2. Identificar o algoritmo de classificação com a maior média de exatidão e a menor quantidade de canais visando a identificação de movimentos de membros inferiores, e

3. Analisar o impacto do número de tarefas na exatidão do treinamento visando a correta classificação/identificação de novas execuções de movimentos de membros inferiores.

2. FUNDAMENTAÇÃO TEÓRICA

Para embasamento teórico e sustentação do projeto seguem os conceitos para melhor compreensão do proposto.

2.1. INTERFACE CÉREBRO-MÁQUINA E ELETROENCEFALOGRAMA

De acordo com Wolpaw *et al.* (2002), um Brain-Computer Interface (BCI) é caracterizado por ser um sistema de comunicação que permite aos indivíduos codificar a atividade cerebral em sinais de EEG. Um BCI consiste em um recurso que inclui tanto hardware quanto software e fornece um canal de comunicação não muscular para pessoas com e sem deficiências (BRUNNER *et al.*, 2011).

Ressalta-se que o hardware BCI registra sinais do cérebro, que podem ser adquiridos de forma invasiva ou não invasiva, usando uma série de componentes do dispositivo. BCI não invasiva é quando se utilizam eletrodos posicionados no escalpo. Aplicações invasivas podem empregar eletrodos diretamente no córtex ou na superfície do cérebro (PANOULAS; HADJILEONTIADIS; PANAS, 2010).

O sinal coletado pelo BCI é conhecido como eletroencefalograma (EEG) e esse sinal é medido pelas correntes geradas pelas sinapses no córtex cerebral. Durante esse processo, ondas magnéticas são geradas pelas correntes e o campo magnético formado pelas ondas são mensuráveis pelos equipamentos denominados eletromiógrafos (SANEI; CHAMBERS, 2007).

Algumas classificações são possíveis. É possível categorizar aplicações de BCIs em quatro categorias: pesquisa básica, clínica/translacional, produtos de consumo e aquelas voltadas a aplicações emergentes (BRUNNER *et al.*, 2011). Também é possível classificá-las em BCIs dependentes e independentes.

O BCI dependente é basicamente um meio de detectar as mensagens carregadas pelo cérebro de forma natural, ou seja, é necessário que o sinal seja carregado de forma natural para que exista um sinal como o EEG. Em contraste, o modelo independente não precisa que os neurônios periféricos carreguem a mensagem. Assim sendo, para um sinal EEG ser lido, apenas é necessário a intenção do usuário em realizar a ação e não necessariamente a realização da ação.

Outra possibilidade é em BCI endógeno e exógeno (CHAN *et al.*, 2015). O primeiro tipo não requer estímulos externos, baseando-se somente em ritmos cerebrais e outros potenciais. O segundo tipo de BCI depende da atividade neuronal evocada por estímulos elétricos. A Figura 1 apresenta uma possibilidade de organização das classes de BCI.

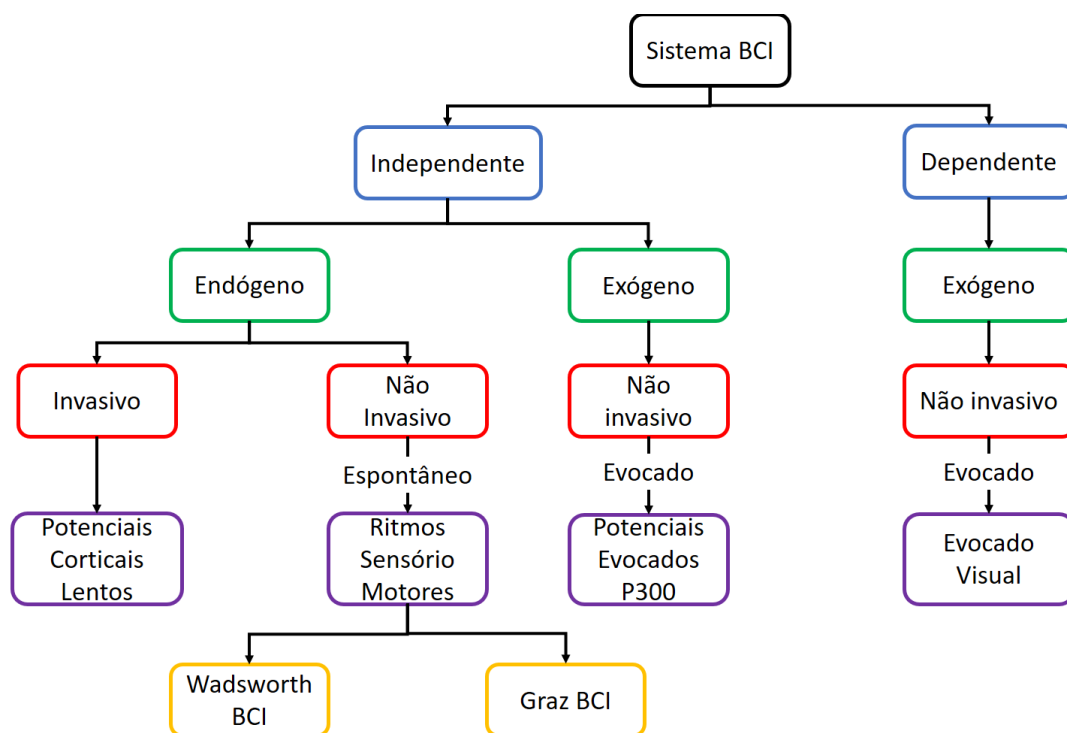


Figura 1 - Taxonomia abstrata do desenvolvimento de BCIs. (Fonte: O autor, 2021)

Ainda segundo Wolpaw *et al.* (2002), a BCI altera sinais eletrofisiológicos de meros reflexos da atividade do sistema nervoso central (SNC) em produtos pretendidos dessa atividade e explica que esses produtos são mensagens e comandos. Basicamente, o BCI gera sinais eletrofisiológicos que o hardware e software traduzem em ações sem a necessidade do movimento muscular.

2.2. REGIÕES DO CÉREBRO

Além disso, é importante destacar que, de forma simplificada, o encéfalo pode ser dividido em três partes principais: cérebro, cerebelo e tronco encefálico. O encéfalo é dividido em dois hemisférios, direito e esquerdo, e tem uma camada externa de massa cinzenta, denominado de córtex cerebral, que por sua vez pode ser

dividido entre os lobos frontal, parietal, temporal e occipital (PURVES *et al.*, 2008). A Figura 2 representa de forma gráfica as divisões.

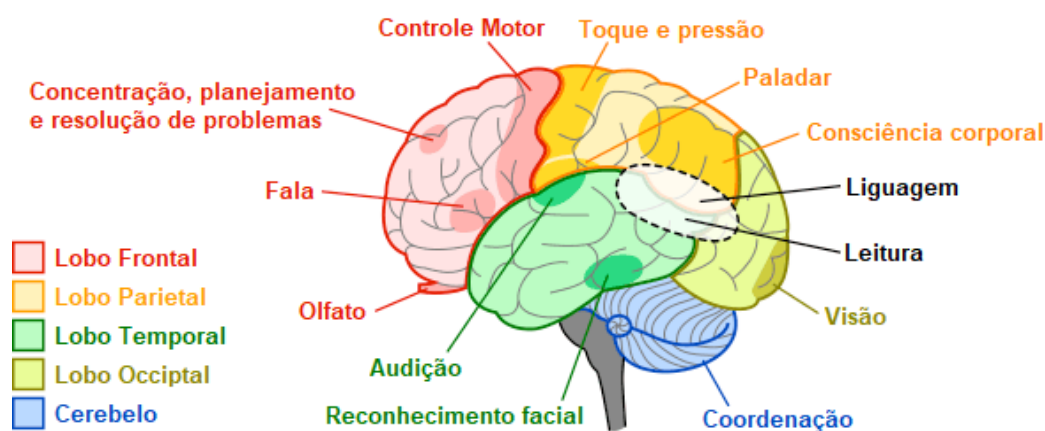


Figura 2 - Regiões do cérebro (Adaptado de SZYMIK, B., 2011)

Para este estudo o foco é no lobo central, sendo essa a maior região do córtex e responsável pelas ações físicas e mentais divididas em cinco áreas: córtex motor primário, córtex motor suplementar, campo ocular suplementar, córtex pré-frontal anterior e a parte opercular. Dentre estas, o córtex motor primário tem a função de iniciar e executar movimento muscular voluntário e o córtex motor suplementar tem a função de planejamento e coordenação dos movimentos (PURVES *et al.*, 2008).

2.3. SISTEMA DE POSICIONAMENTO INTERNACIONAL

Como mencionado anteriormente, durante o procedimento de coleta destes sinais cerebrais, pequenos discos metálicos, chamados eletrodos, são colocados de forma invasiva ou, mais comumente não invasiva, no couro cabeludo do paciente, e os chamados potenciais evocados são os estímulos lidos por eles que geram as informações em forma de sinais eletroencefalográficos (SAZGAR, 2019).

Existem várias maneiras de posicionamento dos eletrodos, mas os mais comuns são o sistema internacional de posicionamento 10-10 ou 10-20, que descrevem o nome de cada eletrodo e o posicionamento correto de cada um deles. O sistema 10-10 é baseado em marcos, definidos por duas linhas de contorno que se

interceptam em 50% da distância total deles, e esse ponto de intercessão é chamado de vértice, sendo este o ponto de referência presente no meio da cabeça para o posicionamento de todos os outros eletrodos (SAZGAR, 2019). A Figura 3 demonstra de forma visual o posicionamento dos eletrodos em ambos os sistemas.

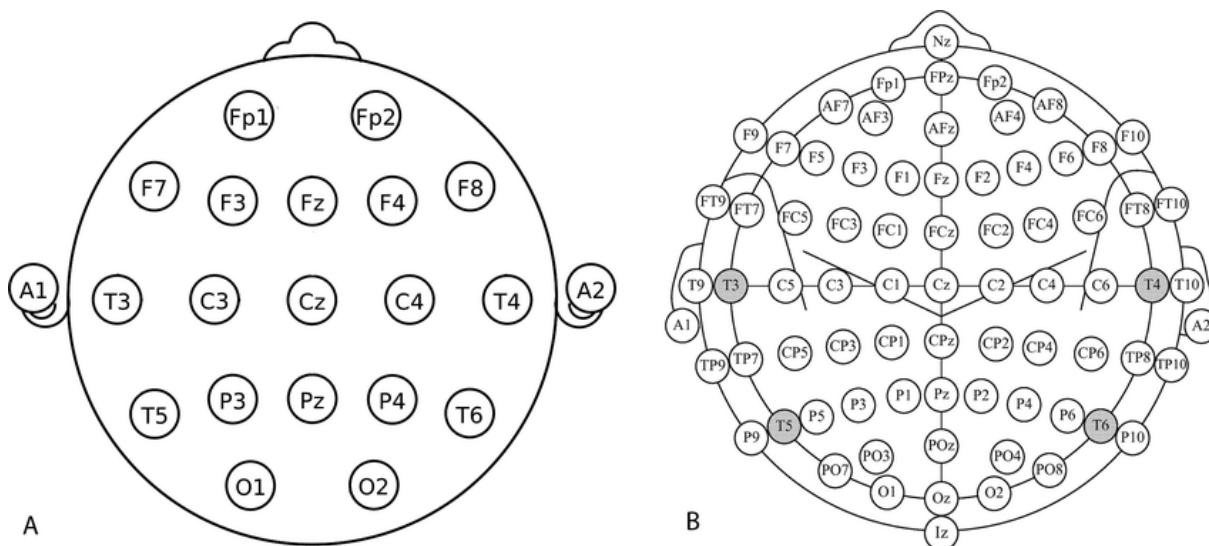


Figura 3 - Sistemas de Posicionamento Internacional. A) Sistema 10-10. B) Sistema 10-20 (MASKELIUNAS et al., 2016).

De modo geral, para o sistema 10-10, o próximo eletrodo é então posicionado à 10% da distância, e isso é repetido para todos os eletrodos. Já no sistema 10-20, essa distância é diminuída para 5%, implicando no surgimento de novos pontos de posicionamento. Para nomear esses pontos, em ambos os sistemas são utilizados letras e números para denominação de cada posição, sendo utilizado desde a frente da cabeça até atrás, as letras FP para pré-frontal, F para frontal, C para a linha central da cabeça, T para a linha temporal, P para a linha parietal e O para a linha occipital, além de números ímpares para o lado esquerdo, pares para o lado direito da cabeça e Z para a linha central (SAZGAR, 2019).

2.4. APRENDIZADO DE MÁQUINA

O aprendizado de máquina ou *machine learning* (ML) é um campo interdisciplinar que se baseia em conceitos de ciência da computação, estatística, ciência cognitiva, engenharia, teoria de otimização e muitas outras disciplinas de matemática e ciências (GHAHRAMANI, 2003). De acordo com Sage (1990 apud

HAYKIN, 2001) um sistema de ML deve ser capaz de realizar tarefas cognitivas similares aos humanos e deve ser capaz de armazenar conhecimento, aplicá-lo em resolução de problemas e adquirir novo conhecimento através da experiência.

Conforme Haykin (2001), “o cérebro é um computador (sistema de processamento de informação) altamente complexo, não-linear e paralelo”. No momento em que nasce um ser humano, o cérebro inicia o processo de desenvolver suas próprias regras, a qual é chamada de “experiência”, sendo assim, permite que o sistema nervoso em desenvolvimento se adapte ao meio-ambiente devido à “plasticidade” do cérebro humano.

Já segundo Mohri, Rostamizadeh e Talwalkar (2018) o ML consiste em realizar um design eficiente e assertivo de algoritmos de previsão. Além disso, ele utiliza as informações previamente disponibilizadas como experiência e utiliza métodos computacionais a partir destas para aprimorar sua performance ou fazer previsões precisas, e devido ao fato de depender dos dados utilizados, ML está inerentemente relacionado às áreas de análise de dados e estatística.

Segundo os autores, ele pode ser aplicado em diversas tarefas como por exemplo: classificação de textos ou documentos, podendo por exemplo detectar spam ou conteúdo inapropriado contido; aplicações de processamento de fala, que inclui reconhecimento de oratória; previsão de diagnóstico médico e muito mais. Além de realização de tarefas de aprendizado amplamente realizadas como classificação, regressão e agrupamento.

Os autores comentam também que os cenários de aprendizagem mais comuns são: (i) aprendizagem supervisionada, em que é informado uma quantidade de dados classificados e prevê os pontos não visualizados; (ii) aprendizagem não supervisionada, em que é informado uma quantidade de dados não classificados e prevê os pontos não visualizados; (iii) aprendizagem semi-supervisionada, em que é informado uma quantidade de dados classificados e não classificados e prevê os pontos não visualizados, e (iv) aprendizagem por reforço, em que para coletar os dados, a máquina deve interagir com o ambiente e isso pode ou não alterá-lo e, após realizar uma ação, ele imediatamente recebe uma recompensa entre outros.

A conjunção entre ML e BCI tem recebido atenção no desenvolvimento de soluções BCI para resolver problemas complexos na área biomédica e se tornou uma ferramenta eficiente para sistemas BCI (MDPI, 2021).

2.5. CLASSIFICADORES

Segundo Holland *et al.* (1999), uma definição de sistemas classificadores pode ser considerada que por meio de um conjunto de detectores, o estado do ambiente é transmitido ao cérebro e as saídas são tratadas como pacotes de informação padronizado, podendo ser chamado de mensagens, e estes, após processamento interno, gerarão novas mensagens encaminhadas aos efetores do sistema, determinando as ações do indivíduo no ambiente.

2.5.1. Support Vector Machine

Segundo Cristianini *et al.* (2000), uma estratégia de aprendizado é o Support Vector Machine (SVM), que consiste em um método poderoso baseado em um algoritmo de aprendizagem de teorias de otimização que implementa conceitos derivados de teorias de aprendizagem estatística treinando hipóteses de espaço de funções lineares em espaços de recurso de alta dimensão. Isto é, o sistema utiliza teorias de aprendizado estatístico e tenta encontrar uma função matemática em um hiperplano ideal.

O método foi originalmente proposto no final dos anos 90 por Cortes and Vapnik (1995). De acordo com (Mammone *et al.*, 2009) a maioria das aplicações encontradas pelos autores são relacionados a bases de dados com relativamente menores números de variáveis, mas também não apresentam restrições para as com conjuntos de dados multivariáveis.

De maneira geral, os autores explicam que o SVM tem por objetivo buscar um “bom” hiperplano de separação em um espaço que pode ser altamente multidimensional de modo eficiente. Ou seja, com base no previamente exposto, o SVM busca um algoritmo que quando exposto graficamente, gera uma função que separa os pontos distribuídos no mapa, tentando otimizar a separação em dois grupos distintos, de maneira a estar apresentado em equidistância dos dois grupos.

2.5.2. Linear Discriminant Analysis

O método de *Linear Discriminant Analysis* (LDA) de acordo com Shashoa et al. (2016) busca uma projeção de vetores discriminantes ótima por meio de transformações linear para assim mapear as características visualizadas no primeiro grupo de informação em outro conjunto de dimensão inferior. Ele realiza reconhecimento de padrões por meio de transformações lineares de modo que a sobreposição entre as classes no espaço de recursos dimensional inferior é mínima.

O LDA, assim como o SVM, utiliza de um algoritmo para determinar onde está o limite de cada classe de resposta, tratando cada amostra como uma amostra em conformidade com uma distribuição normal multidimensional, envolvendo cálculos de vetor médio e a matriz de covariância entre cada classe de dados. Assim, determina-se uma equação de barreira resultando em equações de limite que dependem dos parâmetros da distribuição ajustada (MATLAB & Simulink, 2022).

Se a matriz de covariância for igual para todas as classes de dados, a equação da barreira será linear, e facilitará todo o processo de cálculo e classificação. Caso elas não sejam iguais, será necessário aplicar uma função quadrática e exige mais computação e memória para realizar a classificação (MATLAB & Simulink, 2022).

2.5.3. K-Nearest Neighbor

Kramer (2013) descreve em seu livro que o método *k-Nearest Neighbor* é baseado na classificação dos padrões mais k-próximos, em distâncias euclidianas, no espaço de dados e são conhecidas por terem bom desempenho no caso de grandes conjuntos de dados de baixa dimensão.

Ele continua descrevendo que o kNN é um método de classificação baseado na teoria de que os padrões mais próximos a um padrão-alvo “x”, para o qual buscamos a classificação, fornecem informações úteis sobre a classificação do alvo. Sendo assim, o kNN atribui a etiqueta da classe da maioria dos k padrões mais próximos no espaço de dados.

Também explica que a escolha de k define a localidade do kNN. Para $k = 1$, pequenos “bairros” surgem em regiões onde padrões de diferentes classes estão espalhados. Para tamanhos de “bairro” maiores, como $k = 20$, padrões com etiquetas

na minoria são ignorados. No caso de grandes conjuntos de dados, o kNN precisa procurar os k padrões mais próximos no espaço inteiro, mas já pode fornecer uma boa aproximação com base nos k vizinhos mais próximos em um subconjunto escaneado.

2.5.4. Validação cruzada k -fold

Conforme Berrar (2018) a validação cruzada é um método de reamostragem de dados para avaliar a capacidade de generalização de modelos preditivos e para evitar *overfitting* (quando a máquina se ajusta muito com a base treinada dando uma alta exatidão, mas ao testar novos exemplos, não reconhece muito bem). A amostragem de dados é feita de forma que não ocorra sobreposição de dois conjuntos de teste.

O autor prossegue explicando que na validação cruzada k -fold, o conjunto é particionado em k subconjuntos de tamanhos equivalentes, e é realizado por meio de amostragem aleatória de casos de conjunto sem sobreposição. Então o modelo segue para treinamento em $k-1$ subconjuntos ($D_{train,n}$), em seguida, o modelo é aplicado ao subconjunto restante chamado de conjunto de validação ($D_{val,n}$), sendo assim, o desempenho do classificador é medido a partir deste. A Figura 4 apresenta uma representação de uma validação cruzada 10-fold.

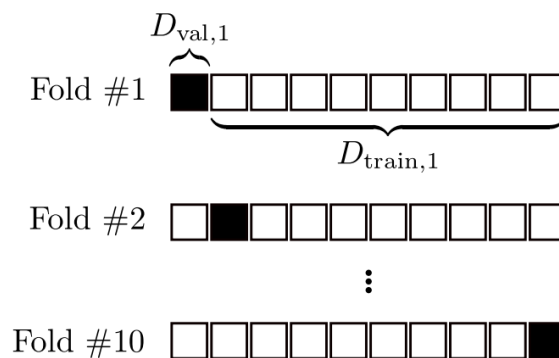


Figura 4 - Representação gráfica de uma validação cruzada 10-fold (Bengio & Grandvalet, 2019)

Como visto na imagem, esse procedimento é repetido até que todos os k subconjuntos tenham servido como conjunto de validação, e nesse caso, k é equivalente a 10. A média de desempenho das k medições dos k conjuntos de validação é o desempenho final após validação cruzada.

2.5.5. Métricas de avaliação

A partir de um treinamento, é gerada uma matriz de confusão. Esta consiste em uma tabela que indica os acertos e erros do modelo treinado, comparando com o resultado esperado (MATLAB & Simulink, 2022). A tabela gerada segue conforme o padrão apresentado pela Tabela 1:

Tabela 1 - Matriz de Confusão (O Autor, 2023)

		Detectada	
		Sim	Não
Real	Sim	Verdadeiro Positivo (VP)	Falso Negativo (FN)
	Não	Falso Positivo (FP)	Verdadeiro Negativo (VN)

Para a Tabela 1, cada linha apresenta o resultado esperado de forma que o classificador tenha acertado ou errado o valor. Se o valor previsto era sim, e ele detectou como sim, então esse será um valor verdadeiro positivo, já se o classificador detectou como não, será um valor falso negativo. Se o valor previsto era não, e ele detectou como sim, então esse será um valor falso positivo, já se o classificador detectou como não, será um valor verdadeiro negativo (MATLAB & Simulink, 2022).

Sendo assim, ao final da classificação de todos os dados, essas informações serão apresentadas como taxas, tendo em vista a proporção da classe com relação ao total de informações, portanto denominados: TPR – *True Positive Rate*, FNR – *True Negative Rate*, FPR- *False Positive Rate* e TNR – *True Negative Rate* (MATLAB & Simulink, 2022).

Além disso, a partir dos resultados obtidos pelo classificador é possível utilizar algumas métricas de avaliação, sendo uma delas a exatidão (MATLAB & Simulink, 2022). A exatidão segue a seguinte equação:

$$\frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

Sendo assim, a exatidão indica a performance geral do modelo, considerando dentre todas as classificações, quantas o modelo classificou corretamente. Esse valor é apresentado de forma percentual para cada base de dados treinada. (MATLAB & Simulink, 2022).

2.6. MATLAB

O MATLAB consiste em um pacote de software que apresenta uma linguagem de programação de alto nível, que permite a expressão de modelos matemáticos baseados em matrizes e uma vasta biblioteca de toolboxes que abrangem desde processamento de sinais e imagens, sistemas de controle, comunicação wireless e computação financeira até robótica, *deep learning*, inteligência artificial e outros (MATLAB & Simulink, 2022).

O MATLAB contém aplicativos para treinar modelos a partir do uso de ML, um deles é o *classification learner*, que simplifica tarefas trabalhosas através de aplicativos para comparação de modelos e possui técnicas populares de classificação, regressão e algoritmos “clusterizados” para aplicação de aprendizados supervisionados e não supervisionados (MATLAB & Simulink, 2022).

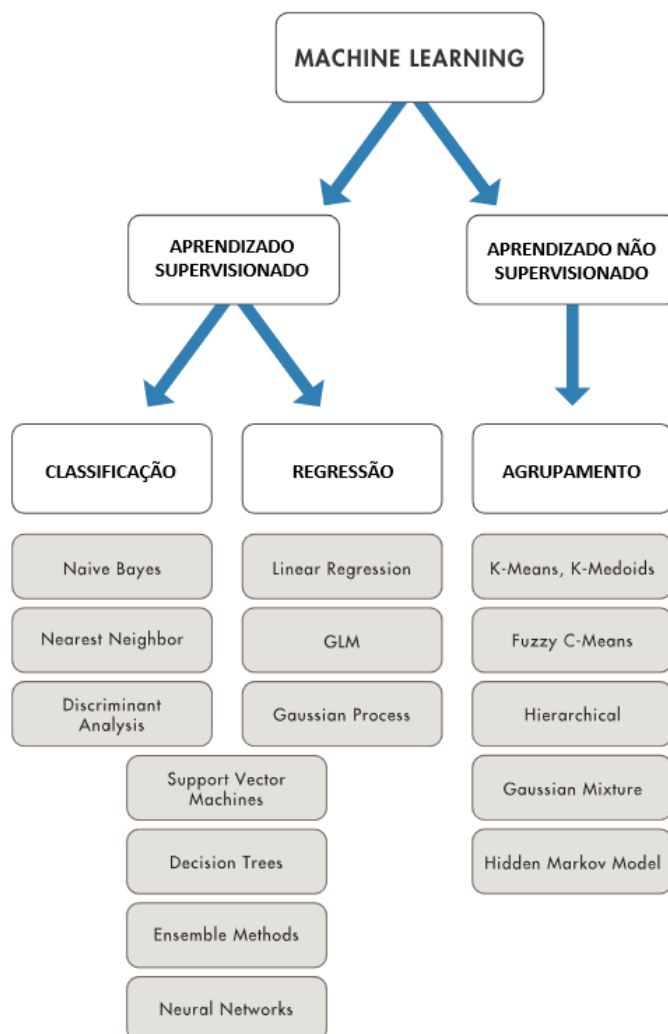


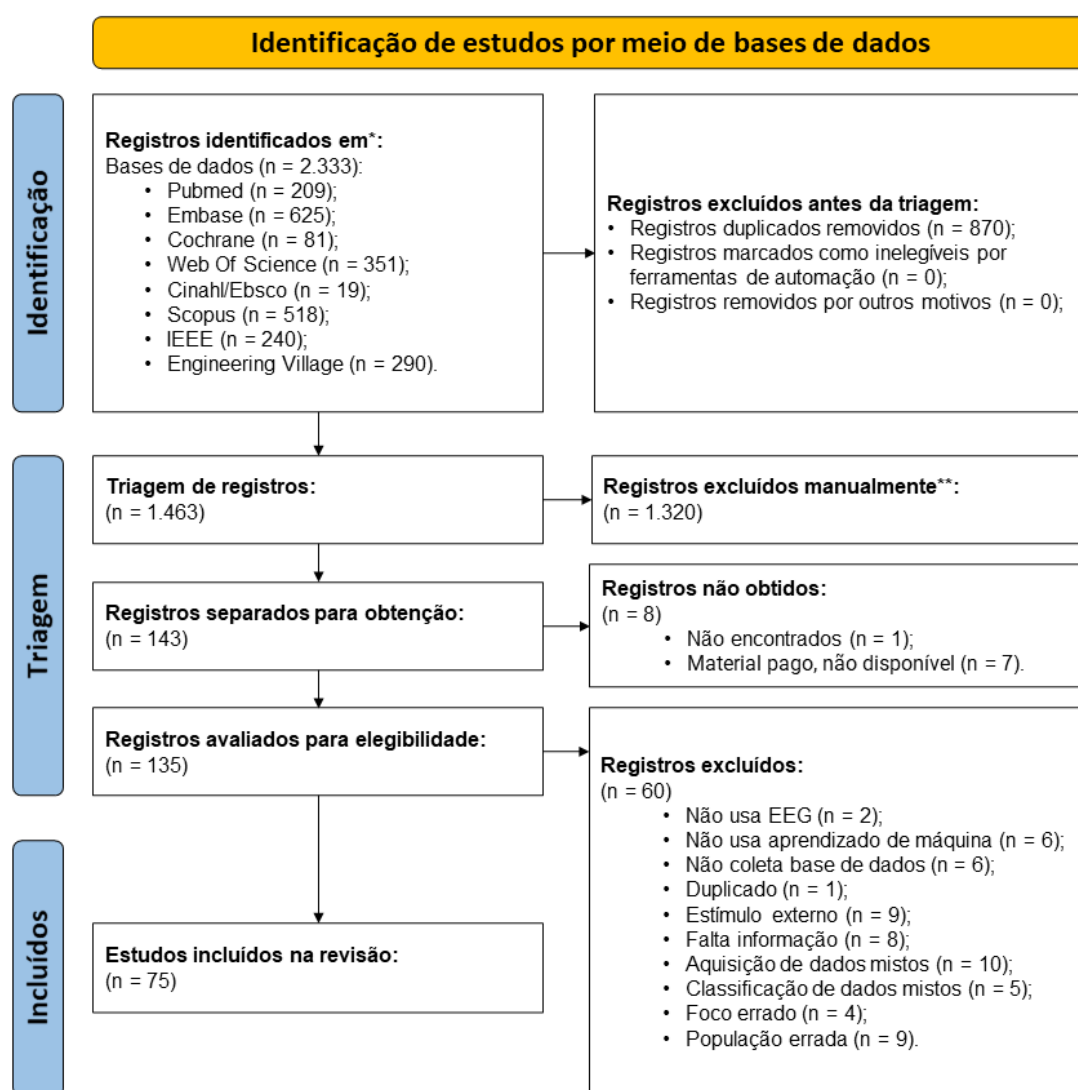
Figura 5 - Métodos de Machine Learning disponíveis no MATLAB (Adaptada de MATLAB & Simulink, 2022).

A Figura 5 apresenta dentro de ambas as classificações de ML, supervisionado e não supervisionado, alguns dos dispositivos possíveis para aplicação de treinamento. O MATLAB também possibilita a utilização ML de treino de modelos *tall array* com grande volume de dados para não sobrecarregar a memória.

O aplicativo focado nesta dissertação foi o *Classification Learner*, que conforme o nome sugere, treina modelos para a classificação de dados. Além disso, permite a utilização dos algoritmos classificadores focados neste trabalho: SVM, LDA e kNN.

3. ESTADO DA ARTE

Para a realização deste trabalho, entendeu-se necessário o levantamento do estado da arte no tema, de modo que se realizou uma revisão de literatura nas bases de dados EMBASE, PUBMED, Cochrane Library, Web of Science, CINAHL, Scopus, IEEE Xplore Digital Library, e Engineering Village, sendo dia 21 de setembro de 2022 a data da última pesquisa realizada.



*Considere, se possível, relatar o número de registros identificados em cada banco de dados ou registro pesquisado (em vez do número total em todos os bancos de dados/registros).

**Se foram usadas ferramentas de automação, indique quantos registros foram excluídos por um humano e quantos foram excluídos por ferramentas de automação.

De: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021;372:n71. doi: 10.1136/bmj.n71. Para mais informações, visite: <http://www.prisma-statement.org/>

Figura 6 - Fluxograma do PRISMA descrevendo as etapas da revisão e artigos selecionados (O Autor, 2023)

A Figura 6 acima apresenta o fluxograma padronizado pela PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*. Page et al., 2021). Nele são destacadas as seções de identificação de artigos, de triagem e de inclusão, descritos a seguir.

Na seção de identificação, as pesquisas realizadas englobaram as seguintes palavras chaves e respectivos sinônimos nas ferramentas de pesquisas das bases de dados: *Brain-Computer Interface* (BCI), Eletroencefalograma (EEG), *Gait*, *Lower Limbs*, *Machine Learning* e *Classification*. Os termos foram buscados em inglês tendo em vista que a maioria dos estudos disponibilizados são escritos em inglês e teria um maior alcance na busca.

Foram localizados 2.333 resultados e, após remoção das duplicatas, 1.463 resultados sobraram, que seguiram para a etapa de triagem passando por revisão de título e resumo. Dessa forma, 1.320 artigos foram excluídos por não atenderem os critérios de inclusão e 135 artigos passaram por revisão do texto completo. Dentre esses, apenas 75 artigos foram mantidos para coleta de dados e análises.

Os critérios de inclusão eram papéis que consistiam em artigos ou conferências com o texto completo disponível, publicados em inglês, que coletavam o sinal EEG a partir de movimentos ou imagética de movimentos de membros inferiores de forma não invasiva em adultos com função motora.

Os artigos que atenderam tais critérios puderam ser excluídos pelos seguintes critérios de exclusão: tinham um método híbrido de coleta de sinal (por exemplo: coletava sinais de EEG em conjunto com EMG e EOG), os sinais EEG eram coletados para outros tipos de movimentos em conjunto com os incluídos sem a alteração da seleção de canais, o foco era em movimentos involuntários ou era coletado movimentos com estímulo externo (aplicação de estímulo externo para gerar o EEG) e intervenções realizadas em animais.

Além destes, também foram excluídos pelos seguintes motivos: quando sinais EOG eram coletados propositalmente, quando a classificação era realizada com mais de um tipo de sinal coletado, por exemplo EEG+EOG, e quando eram revisões ou capítulos de livros.

A Tabela 2 a seguir apresenta o resultado obtido a partir da realização da revisão de literatura. Consiste em uma tabela que apresenta as melhores configurações de canais encontradas durante o estudo e serviu de base para direcionar os treinamentos realizados nessa dissertação.

Como critério de exclusão para elaboração da Tabela 2, todos os artigos que tiveram exatidão menor que 80% não foram considerados para as validações, e caso alguma quantidade de canais estivesse repetida, o estudo que apresentava maior exatidão foi selecionado para seguir com exceção de um canal que, pelo fato de o canal CZ ter sido o mais utilizado entre todos os estudos, entendeu-se por necessário comparar com a classificação do canal FP1.

Ainda, é importante destacar que, para a metodologia de reamostragem de dados, fez-se uma busca na base de dados e o método mais utilizado foi o de validação cruzada, sendo a validação cruzada 10-fold a mais aplicada entre todos os estudos encontrados.

Tabela 2 - Quantidade e seleção de canais escolhidos (O Autor, 2023).

Autor	Título	Exatidão	Seleção de canais	Total de Canais
Kuramoto, N; Ito, S; Sato, K; Fujisawa, S	Trigger pattern detection method for assisting in ambulation rehabilitation based on EEG analysis	91,20%	FP1	1
Rodríguez-Ugarte, M; Ângulo-Sherman, I N; Ianez, E; Ortiz, M; Azorin, J M	Preliminary study of pedaling motor imagery classification based on EEG signals	85,00%	CZ	1
Rodríguez-Ugarte, M; Ianez, E; Ortiz, M; Azorin, J M	Personalized Offline and Pseudo-Online BCI Models to Detect Pedaling Intent	93,80%	FC2, FC6, C2, C4	4
Rodríguez-Ugarte, M; Ianez, E; Ortiz, M; Azorin, J M	Personalized Offline and Pseudo-Online BCI Models to Detect Pedaling Intent	81,80%	C3, C1, CZ, C2, C4	5
Somadder, R; Saha, D K	Frequency Domain CSP for Foot Motor Imagery Classification Using SVM for BCI Application	84,85%	F3, FZ, F4, C3, CZ, C4	6
Li, H; Zhang, X; Lu, Z; Dong, R; Sun, Q	RP-based Voluntary Movement Intention Detection of Lower limb using CNN	84,20%	FC1, FCZ, FC2, C1, CZ, C2, CP1, CP2	8
Soriano-Segura, P; Iáñez, E; Quiles, V; Ferrero, L; Ortiz, M; Azorín, J M	Selection of Spatial, Temporal and Frequency Features to Detect Direction Changes During Gait	88,10%	FC1, FCZ, FC2, C1, CZ, C2, CP1, CPZ, CP2	9
Rodríguez-Ugarte, M; Ianez, E; Ortiz, M; Azorin, J M	Personalized Offline and Pseudo-Online BCI Models to Detect Pedaling Intent	82,80%	FZ, FC5, FC1, FC2, FC6, C3, C1, CZ, C2, C4	10
Malik, A N; Iqbal, J; Tiwana, M I	EEG signals classification and determination of optimal feature-classifier combination for predicting the movement intent of lower limb	81,48%	AF3, F&, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4	14
Planelles, D; Hortal, E; Costa, A; Iáñez, E; Azorín, J M	First steps in the development of an EEG-based system to detect intention of gait initiation	80,69%	FZ, FC5, FC1, FCZ, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P3, PZ, P4	16
Triana-Guzman, Nayid; Orjuela-Cañon, Alvaro D; Jutinico, Andres L; Mendoza-Montoya, Omar; Antelis, Javier M	Decoding EEG rhythms offline and online during motor imagery for standing and sitting based on a brain-computer interface.	96,56%	F3, FZ, F4, FC5, FC1, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P3, PZ, P4	17
Delisle-Rodriguez, Denis; Cardoso, Vivianne; Gurve, Dharmendra; Loterio, Flavia; Alejandra Romero-Laiseca, Maria; Krishnan, Sridhar; Bastos-Filho, Teodiano	System based on subject-specific bands to recognize pedaling motor imagery: towards a BCI for lower-limb rehabilitation	96,43%	FZ, FC3, FC1, FCZ, FC2, FC4, C5, C3, C1, CZ, C2, C4, C6, CP3, CP1, CPZ, CP2, CP4, PZ	19
Soriano-Segura, P; Iáñez, E; Quiles, V; Ferrero, L; Ortiz, M; Azorín, J M	Selection of Spatial, Temporal and Frequency Features to Detect Direction Changes During Gait	93,05%	FC5, FC3, FC1, FCZ, FC2, FC4, FC6, C1, CZ, C2, CP5, CP3 CP1, CPZ, CP2, CP4, CP6, P1, PZ, P2	20

A partir da revisão sistemática, concluiu-se que os classificadores mais utilizados eram SVM, LDA e kNN, que os canais mais utilizados eram C3, CZ e C4 e que era possível reconhecer movimentos de membros inferiores a partir de apenas um canal. A Figura 7 abaixo foi elaborada para representar um mapa de calor em que vermelho representaria os canais mais usados, e gradativamente a cor chegaria até o verde, representando os canais menos usados.

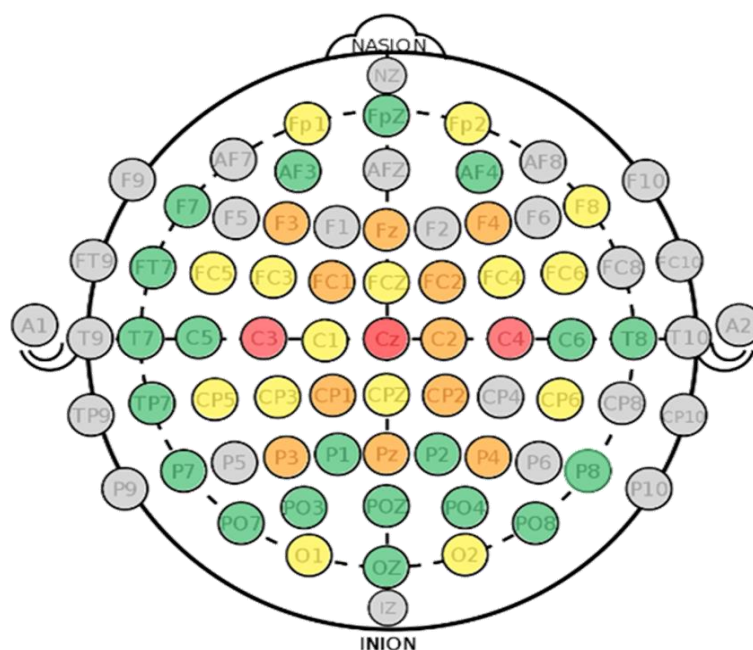


Figura 7 - Mapa de calor dos eletrodos mais utilizados encontrados na literatura (O Autor, 2023)

Além disso, os resultados relacionando redução de canais para os algoritmos mais utilizados não foram conclusivos para afirmar que diminuir a quantidade de canais aumenta ou diminui a exatidão do classificador, justificando mais uma vez a necessidade deste estudo.

Em relação aos estudos que envolviam classificação pelo algoritmo SVM, o trabalho que encontrou a maior exatidão foi o estudo de Rodríguez-Ugarte, Ianez, Ortiz e Azorin (2017). A exatidão encontrada foi de $86,70\% \pm 8,16\%$ com apenas 5 canais.

O estudo consistia em coletar os sinais de EEG de 5 pacientes durante a movimentação de pedalada com auxílio de uma bicicleta de fisioterapia. O sinal coletado passou por cinco métodos de filtragem e extração de características e foi

classificado com o software MATLAB de maneira offline e pseudo-*offline*. A coleta do sinal foi realizada com todos os 32 canais disponíveis no BCI posicionados de acordo com o sistema internacional de posicionamento 10-10.

Ainda, os autores aplicaram um método de seleção de canais foi aplicado e oito configurações diferentes foram testadas. O melhor resultado obtido pela pesquisa foi com a combinação de 5 canais, sendo eles FC1, FC2, C3, CZ e C4.

Já para os estudos que classificaram o sinal por meio do algoritmo LDA, o estudo que obteve a maior exatidão foi o de Triana-Guzman *et al.* (2022). A exatidão encontrada atingiu $96,56\% \pm 0,83\%$ com 17 canais, sendo eles F3, Fz, F4, FC5, FC1, FC2, FC6, C3, CZ, C4, CP5, CP1, CP2, CP6, P3, PZ e P4.

Neste estudo foi coletado o sinal EEG de 32 pacientes por meio de 17 eletrodos e os movimentos realizados para a coleta eram sentar-se e levantar-se de uma cadeira. Também foram realizadas rodadas de execução motora e de imagética motora separadamente. As classificações de levantar-se e sentar-se foram realizadas separadamente, e o melhor resultado obtido foi de sentar-se.

Por fim, o estudo que obteve a maior exatidão com o algoritmo kNN foi o de Soriano-Segura, Iáñez, Ferrero, Ortiz e Azorín (2020). A exatidão encontrada foi de $95,10\% \pm 3,60\%$ com 20 canais, sendo FC5, FC3, FC1, FCZ, FC2, FC4, FC6, C1, CZ, C2, CP5, CP3, CP1, CPZ, CP2, CP4, CP6, P1, PZ, P2.

O estudo foi realizado com 4 pacientes, e a movimentação realizada no estudo era o ato de caminhar, que consistia em um caminhar contínuo por no mínimo 5 s e então o paciente poderia mudar a sua direção, para a direita ou a esquerda, de acordo com sua vontade. Esse estudo realizou a coleta dos sinais com 32 canais, e os modelos de 20 e 9 canais foram estudados, sendo que o de 20 obteve a maior exatidão do estudo.

Três artigos estudaram apenas um canal: Decoding Movement-Related Cortical Potentials Based on Subject-Dependent and Section-Wise Spectral Filtering (JEONG, KWAK, GUAN, LEE, 2020), Preliminary study of pedaling motor imagery classification based on EEG signals (RODRÍGUEZ-UGARTE, ANGULO-SHERMAN, IANEZ, ORTIZ E AZORIN, 2017), com exatidão de até $85,20\% \pm 12,64\%$ e Trigger pattern detection

method for assisting in ambulation rehabilitation based on EEG analysis (KURAMOTO, ITO, SATO, FUJISAWA, 2012).

Os estudos de Jeong *et al.* (2020), com exatidão de até $86,00\% \pm 9,00\%$ para SLDA, e de Rodríguez-Ugarte *et al.* (2017), com exatidão de até $85,20\% \pm 12,64\%$ para LDA e $84,60\% \pm 13,16\%$ para SVM, estudaram o canal CZ. Já o de Kuramoto *et al.* (2012) estudou a classificação do canal FP1, obtendo a maior exatidão entre os três ($90,00\% \pm 20,00\%$ para o classificador SVM).

A listagem de todos os estudos encontrados para a realização desse estado da arte encontra-se disponibilizada nos apêndices deste trabalho, além de ser recomendada a leitura de todos os artigos apresentados e referenciados.

4. MATERIAIS E MÉTODOS

Neste capítulo, descreve-se a metodologia aplicada e os materiais utilizados para realização deste estudo.

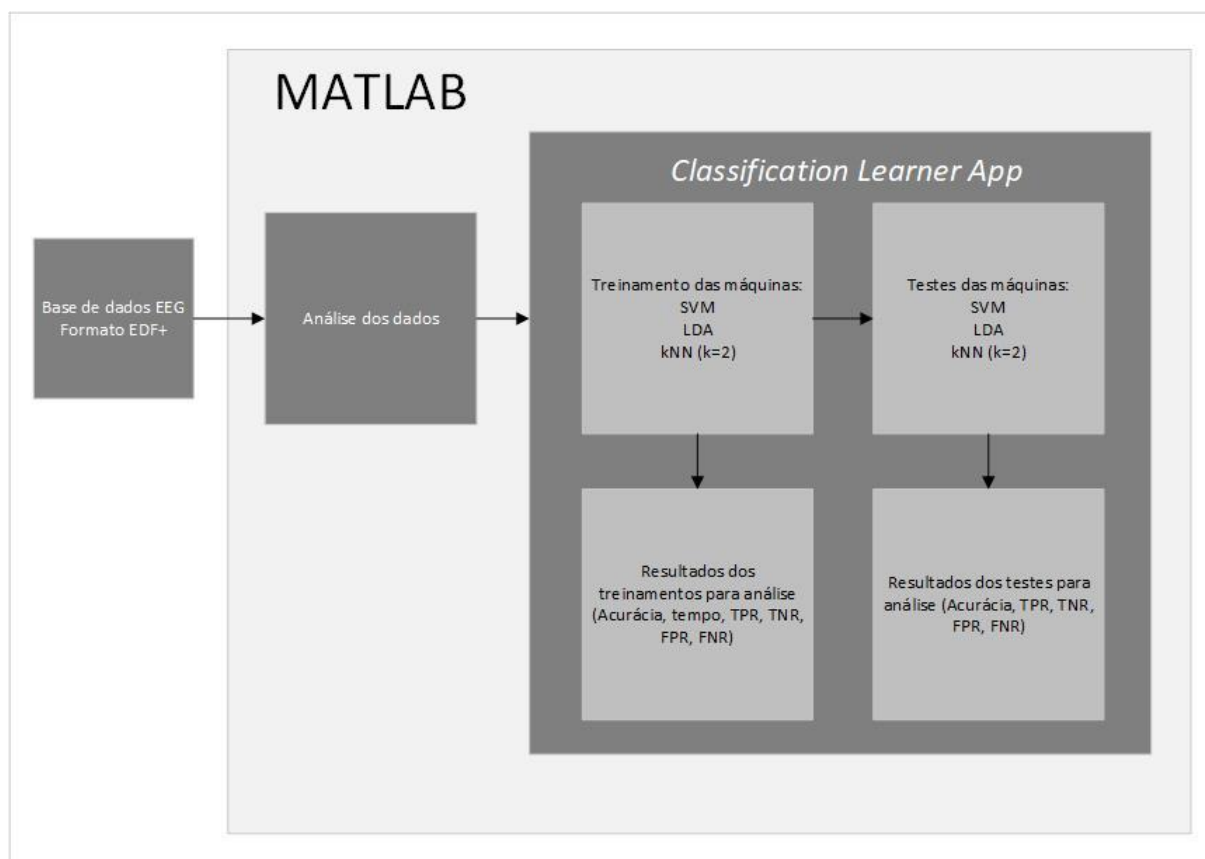


Figura 8 - Diagrama resumindo a metodologia utilizada (O Autor, 2023).

A Figura 8 apresenta de forma resumida a metodologia aplicada. Inicialmente foi definida uma base de dados previamente coletada por outros autores. Em sequência, por meio do software MATLAB, foi realizada a análise de dados, e então, os treinamentos e testes das máquinas. O detalhamento de cada etapa será apresentado nas sessões a seguir.

4.1. BASE DE DADOS

Para a realização deste, foi necessário buscar entre o acervo disponibilizado de outros trabalhos de coleta de sinais EEG por meio de BCI para registro de execução de movimentos inferiores, uma base de dados confiável e renomada para

realização dos treinamentos para validação de hipóteses descritas na seção 1.3.1 Objetivos específicos.

O banco de dados escolhido foi o realizado pelo estudo de Schalk *et al.* (2004), disponibilizado no PhysioBank (Goldberger *et al.*, 2000), que consiste em um banco de dados disponível em formato EDF+ com mais de 1500 gravações de sinais EEG de aproximadamente 1 e 2 min de 64 canais posicionados utilizando o sistema internacional 10-10, com uma taxa de amostragem de 160Hz, obtidos a partir de 109 voluntários.

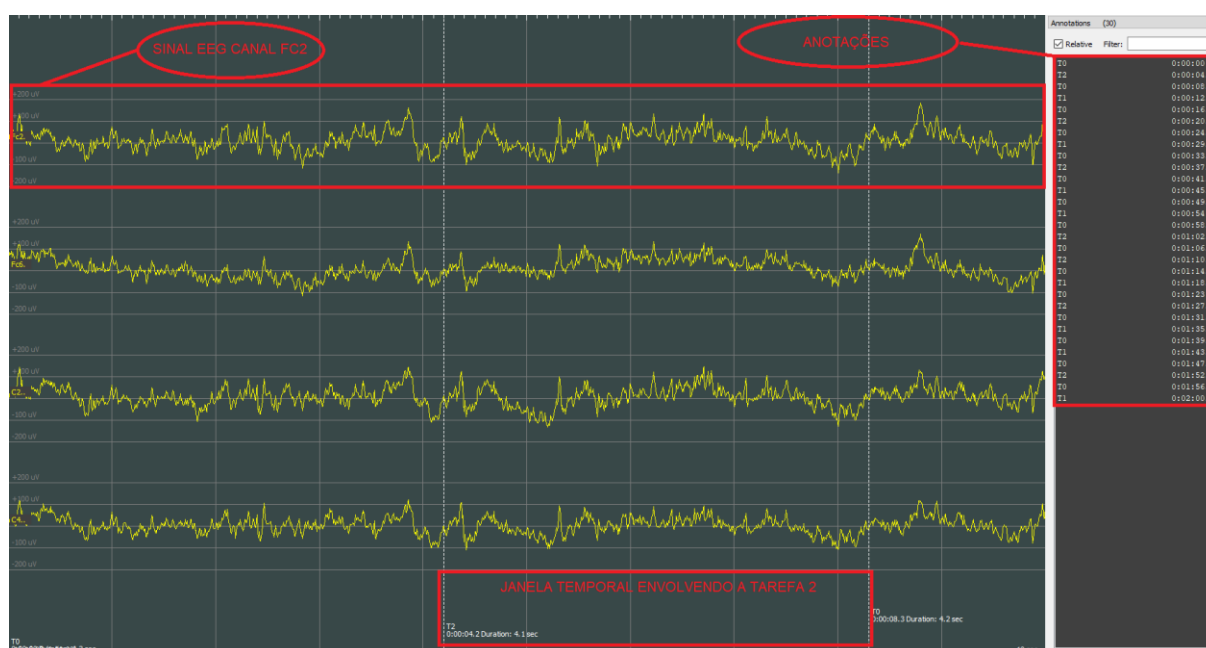


Figura 9 - Representação gráfica do sinal EEG (O Autor, 2023)

A Figura 9, representa a forma de onda de um sinal EEG (4 canais), o qual é separada por marcações (*time-stamps*) durante todo o período da coleta, indicado pelos autores da coleta a tarefa ao qual foi realizada naquela marcação e naquele período de tempo (janela temporal). Ao lado direito é possível ver as anotações feitas pelos autores indicando o início da tarefa executada.

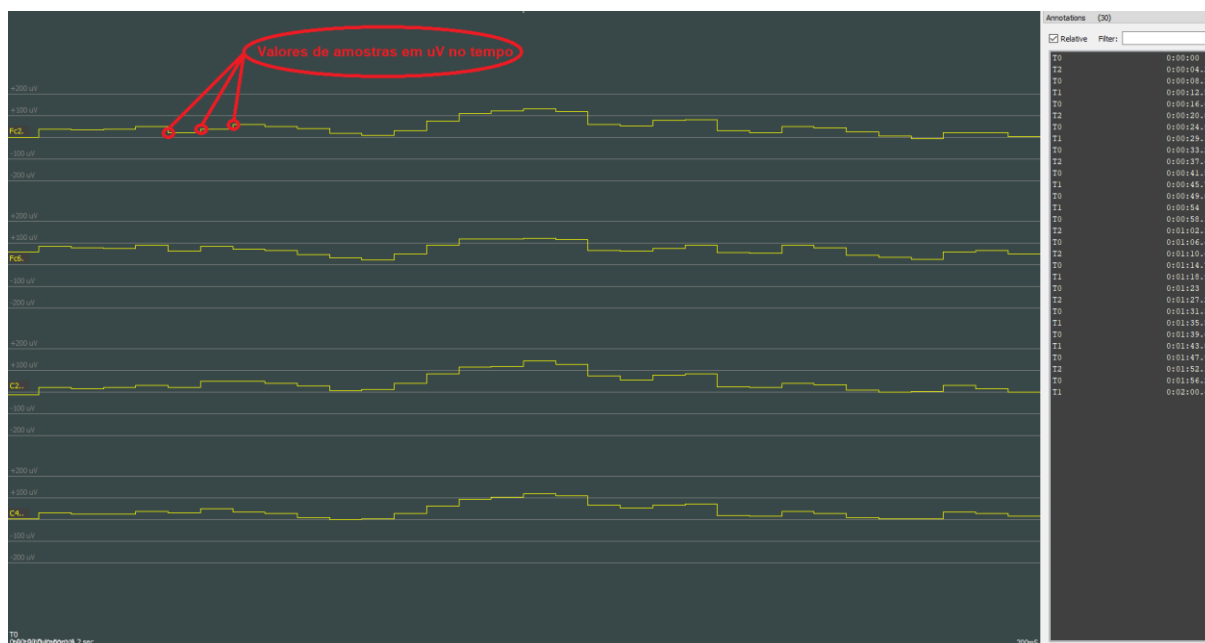


Figura 10 - Representação gráfica de sinal EEG enfatizando os valores das amostras coletadas (O Autor, 2023)

A Figura 10 mostra um *zoom* dentro do sinal EEG mostrado na Figura 9 em que é possível verificar que o sinal é composto de várias amostras temporais (160Hz, 0,00625 amostras por segundo), onde cada amostra é representada por um valor em μV (*microvolt*), o qual representa o valor medido naquele instante.

O protocolo seguido pelo experimento consiste em execução e imagética motora, em que cada voluntário realizou 14 execuções a serem coletadas, numeradas de 1 a 14, variando entre movimentos de mãos e pés. As execuções 5, 9 e 13 foram selecionadas para este estudo, tendo em vista o foco na movimentação de membros inferiores.

Além disso, as execuções (de aproximadamente 2 min) foram separadas em 3 tarefas, T0, T1 e T2, que consistiam em baseline, movimento dos punhos e movimento dos pés respectivamente. Essas tarefas foram executadas aleatoriamente, totalizando 30 tarefas por execução. Tendo em vista que cada voluntário realizou 3 execuções, foram coletadas 90 tarefas por indivíduo que serão utilizadas como exemplos de treinamento para as MLs.

Sendo assim, as tarefas T0 e T1 foram considerados como a classe de não execução de movimentos em membros inferiores (0) e T2 como a execução do movimento (1).

A partir desses 109 voluntários, foi possível excluir os dados coletados de 4 pacientes, tendo em vista que os dados obtidos da coleta destes estavam fora do padrão, portanto esse trabalho foi realizado com os 105 pacientes restantes.

Os voluntários escolhidos para a realização do trabalho foram classificados de forma coletiva separado em dois grupos: um com todos os 105 voluntários e outro com 15 voluntários. Além disso, foram também realizadas classificações individuais dentro do grupo de 15 voluntários.

Reforça-se que no momento de procura da base de dados mais compatível para o intuito deste projeto apenas a desenvolvida por Schalk *et al.* (2004) enquadrava-se nos quesitos necessários, sendo assim, nenhuma outra base teria as informações necessárias para a realização desse.

4.2. HARDWARE E SOFTWARE

Com a base de dados definida, foi possível iniciar os treinamentos de ML para análise de resultados. Para este estudo foi utilizado o software MATLAB R2022b, update 2 (9.13.0.2105380) para o Windows 10 64-bits e versão de atualização em 26 de outubro de 2022.

Para padronizar os treinamentos, foi decidido utilizar o padrão de treinamento fornecido pelo próprio MATLAB para os algoritmos SVM, LDA e kNN, sendo que não foi utilizado nenhum otimizador. Além disso, foi utilizado *Misclassification cost* padrão. Não foram implementados nenhuma seleção de característica e nem extratores de características. No caso do SVM (Figura 11) foi utilizado o *Kernel* Linear, 1 nível de restrição de caixa, modo da escala de Kernel automática e dados padronizados. No LDA (Figura 12) foi utilizada estrutura de covariância completa e para o kNN (Figura 13), k=2 vizinhos, distância euclidiana com peso igual e dados padronizados. Todas as características padrão do MATLAB, excluindo o k=2 vizinhos do classificador kNN que foi decidido utilizar através da revisão de literatura já mencionada anteriormente.

Model 3: SVM
Status: Draft

▼ **Model Hyperparameters**

Kernel function:

Box constraint level:

Kernel scale mode:

Manual kernel scale:

Multiclass method:

Standardize data:

[Read more about SVM model options](#)

► Feature Selection: 656/656 individual features selected

► PCA: Disabled

▼ Misclassification Costs: Default

		Predicted Class	
		0	1
True Class	0	0	1
	1	1	0

 To modify misclassification costs, use the Costs button.

► **Optimizer: Not applicable**

Data set: indarray Observations: 90 Size: 609 kB Predictors: 656 Response: column_657 Response Classes: 2 Validation: 10-fold cross-validation

Figura 11 - Configuração padrão do SVM (O Autor, 2023)

Model 2: Linear Discriminant
Status: Draft

▼ **Model Hyperparameters**

Covariance structure:

[Read more about Discriminant model options](#)

► Feature Selection: 656/656 individual features selected

► PCA: Disabled

▼ Misclassification Costs: Default

		Predicted Class	
		0	1
True Class	0	0	1
	1	1	0

 To modify misclassification costs, use the Costs button.

▼ **Optimizer: Not applicable**

Optimizer options apply only to optimizable models.

Data set: indarray Observations: 90 Size: 609 kB Predictors: 656 Response: column_657 Response Classes: 2 Validation: 10-fold cross-validation

Figura 12 - Configuração padrão do LDA (O Autor, 2023)

Model 4: KNN
Status: Draft

▼ **Model Hyperparameters**

Number of neighbors

Distance metric

Distance weight

Standardize data

[Read more about KNN model options](#)

► Feature Selection: 656/656 individual features selected

► PCA: Disabled

▼ Misclassification Costs: Default

		Predicted Class	
		0	1
True Class	0	0	1
	1	1	0

 To modify misclassification costs, use the Costs button.

► Optimizer: Not applicable

Data set: indarray Observations: 90 Size: 609 kB Predictors: 656 Response: column_657 Response Classes: 2 Validation: 10-fold cross-validation

Figura 13 - Configuração padrão do kNN (O Autor, 2023)

Apesar de existirem diversos algoritmos classificadores, como *Random Forest* e *XBoost*, que estão em ascensão na área, eles não foram incluídos nesse estudo devido ao fato de que, na revisão da literatura, eles não terem sido encontrados com os critérios de classificação selecionados com uma exatidão relevante comparados aos SVM, KNN e LDA escolhidos. Todavia, os mesmos poderiam ter sido utilizados.

A especificação técnica do computador utilizado foi um sistema operacional de 64-bits e processador x64-based. Processador Intel® Core™ i5-7600K CPU @ 3.8GHz, memória RAM de 16GB com velocidade de 2133MHz. A edição do Windows foi a Windows 10 Pro com OS build 19045.2486. Placa de vídeo NVIDIA GeForce GTX 1070 com memória GPU dedicada de 8.0GB GDDR5 e dois HDs de 512GB M.2 NVMe até 3200 MB/s de escrita.

4.3. TREINAMENTOS

A partir da base de dados definida, foi possível com a utilização do MATLAB a programação de um algoritmo que realizou a análise das informações, separando informações dos canais desejados (1 Cz, 1 Fp1, 4, 5, 6, 8, 9, 10, 14, 16, 17, 19 e 20, com as mesmas configurações dos canais que obtiveram os melhores resultados a

partir da revisão de literatura, vide Tabela 2) e identificando a seção de cada tarefa (T0, T1 e T2) de cada indivíduo.

Para a análise dos dados, foi primeiramente realizada uma análise individual por canal, que seguiu conforme descrito abaixo, para cada quantidade de canal selecionada:

- Foi escolhida uma das quantidades e configurações de canais mencionados anteriormente;
- A partir de 64 canais utilizados na coleta da base, foram separados somente os canais do ponto anterior;
- Utilizando as anotações fornecidas pelos autores da base, foram identificados os trechos correspondentes a cada tarefa realizada durante a execução (T0, T1 e T2);
- Com uma amostragem de 160Hz, é extraído um valor (característica do sinal) a cada 0,00625 s; com tempo fornecido por cada tarefa de 4,1 s, tem-se 656 características por canal de cada tarefa;
- Para configurações de canais com mais de 1 canal, foram colocados cada exemplo de teste (tarefa) em um arranjo sequencial, denominado *array* (ex. `vet[656*NumCanais]`);
- Na última coluna do *array* foi etiquetado um valor correspondente a tarefa realizada (ex: `vet[656*NumCanais+1] = 0` ou `1`, lembrando que 0 é a classe de não execução de movimentos em membros inferiores e 1 é a execução do movimento);
- Sendo que cada voluntário realizou 3 execuções de aproximadamente 2 min, em cada execução foram realizadas 30 tarefas (exemplos de teste), totalizou-se 90 exemplos de teste, com `656*NumCanais` número de características, mais 1 da etiqueta da tarefa, para serem treinadas pelas ML;
- Todos os exemplos de teste foram colocados em uma matriz, onde cada linha são os exemplos de teste e as colunas são as características do sinal EEG coletado e a última coluna corresponde à identificação de qual tarefa foi realizada (ex. `mat[exemplos de teste][656*NumCanais+1]`).

Para a análise coletiva, seguiu-se:

- A partir das matrizes de cada indivíduo, foram concatenadas verticalmente, gerando outra matriz maior (ex. `matcoletivo[exemplos de teste*quantidade de voluntários][656*NumCanais+1]`).

Alguns exemplos das matrizes geradas estão descritos a seguir:

- Individual com 4 canais e 90 exemplos de teste:
 - `mat[90][2625]`;
- Coletivo de 15 indivíduos com 4 canais e 90 exemplos de teste:
 - `matcoletivo15[1350][2625]`;
- Coletivo de 105 indivíduos com 4 canais e 90 exemplos de teste:
 - `matcoletivo105[9450][2625]`.

Além disso, foi definida a aplicação de técnicas de ML baseada nos algoritmos SVM, LDA e kNN, definidos a partir da revisão da literatura realizada, descrita na Tabela 2. Assim, pode-se realizar as atividades de classificação seguindo o seguinte fluxo:

- Utilizar as matrizes obtidas através da análise dos dados;
- Utilizar o aplicativo “Classification Learner” dentro do software MATLAB;
- Realizar as classificações utilizando os três classificadores (SVM, LDA e kNN), da seguinte maneira:
 - Individual com os 30 primeiros exemplos de teste do grupo de 15 pessoas mencionado no tópico 4.1;
 - Individual com todos os 90 exemplos de teste do grupo de 15 pessoas;
 - Coletivo do grupo de 15 voluntários para todos os 90 exemplos de teste de cada indivíduo;
 - Coletivo do grupo de 105 voluntários para todos os 90 exemplos de teste de cada indivíduo.

Depois do treinamento dos classificadores, foram obtidos os valores de exatidão e, a partir disso, foram realizadas as seguintes tarefas:

- Tarefa 1. Identificar e selecionar a melhor configuração de canal, visando a redução de canais e o melhor classificador a partir das médias de ac de todos os treinamentos;
- Tarefa 2. A partir do melhor classificador e configuração de canais, escolher o indivíduo com maior exatidão;
- Tarefa 3. A partir do indivíduo definido na tarefa 2, foi utilizada a máquina treinada de 30 tarefas e foram testadas as outras 60 tarefas (90 tarefas totais do indivíduo desconsiderando as 30 já treinadas), para entender como se comporta a eficiência da máquina, mesmo englobando tarefas não treinadas;
- Tarefa 4. A partir do indivíduo com maior exatidão, treinar uma máquina de 60 tarefas e testar as outras 30 tarefas (desconsiderar as 60 tarefas das 90 que foram utilizadas para treinar a máquina), para verificar como se comporta a eficiência da máquina, englobando tarefas não treinadas;
- Tarefa 5. Reduzir a quantidade de exemplos de zeros (T0 e T1) para o melhor indivíduo, na melhor configuração de canais selecionado na tarefa 1 e para todos os classificadores (SVM, LDA e kNN), para verificar como se comportam a exatidão e a medida de “True-positive rate” (TPR), “True-negative rate” (TNR), “False-positive rate” (FPR) e “False-negative rate” (FNR).

5. RESULTADOS

Após a realização da análise dos dados, seguiu-se com o início dos treinamentos considerando todas as iterações entre os três classificadores (SVM, LDA e kNN), as 13 configurações de canais, 15 sujeitos classificados individualmente ou coletivamente, 105 indivíduos classificados coletivamente e 30, 60 ou 90 tarefas, previamente descritas na seção 4.3 TREINAMENTOS REALIZADOS, sendo assim, 1.407 diferentes treinamentos foram realizados.

Com relação aos primeiros testes realizados, considerando 30 exemplos de execução, 15 sujeitos individualmente, os três classificadores e todas as 13 configurações de canais, os resultados médios obtidos são apresentados na Tabela 3.

Tabela 3 - Resultados médios obtidos a partir dos treinamentos realizados para 30 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais e 15 sujeitos individuais (O Autor, 2023).

ML	Média de Acc (%)	Média de TNR (%)	Média de TPR (%)	Média de FPR (%)	Média de FNR (%)
SVM	74,99±4,26	98,18±2,94	5,58±16,12	1,82±2,94	94,42±16,12
KNN	73,5±6,84	92,06±9,39	18,6±19,46	7,99±9,43	81,4±19,46
LDA	73,42±9,91	85,91±9,49	36,14±25,05	14,09±9,49	63,86±25,05

Como é possível verificar na Tabela 3, os valores estão apresentados em ordem decrescente. Assim sendo, o classificador SVM teve o melhor desempenho médio entre os três, teve um TNR e TPR médios equivalente a 98,18%±2,94% e 5,58%±16,12%, respectivamente.

Com relação aos primeiros testes realizados, considerando 90 exemplos de execução, 15 sujeitos classificados individualmente, os três classificadores e todas as 13 configurações de canais, os resultados médios obtidos são apresentados na Tabela 4.

Tabela 4 - Resultados médios obtidos a partir dos treinamentos realizados para 90 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais e 15 sujeitos individuais (O Autor, 2023).

ML	Média de Acc (%)	Média de TNR (%)	Média de TPR (%)	Média de FPR (%)	Média de FNR (%)
SVM	77,54±4,75	96,97±3,12	20,03±21,84	3,03±3,12	79,97±21,84
KNN	74,18±5,83	91,93±7,55	22±17,13	8,07±7,55	78±17,13
LDA	74,06±7,55	84,01±7,17	44,8±17,14	16,04±7,13	55,2±17,14

Na Tabela 4 são apresentados os valores também em ordem decrescente. Dessa forma, o classificador SVM teve o melhor desempenho médio entre os três, teve um TNR e TPR médios equivalente a $96,97\% \pm 3,12\%$ e $20,03\% \pm 21,84\%$, respectivamente.

Por fim, considerando 90 exemplos de execução, os três classificadores e todas as 13 configurações de canais e os coletivos de 15 e 105 voluntários para a classificação, os resultados médios obtidos são apresentados na Tabela 5.

Tabela 5 - Resultados médios obtidos a partir dos treinamentos realizados para 90 exemplos de tarefa, considerando os três classificadores, 13 configurações de canais, 15 sujeitos coletivos e 105 sujeitos coletivos (O Autor, 2023).

ML	Média de Acc (%)	Média de TNR (%)	Média de TPR (%)	Média de FPR (%)	Média de FNR (%)
SVM	76,03 $\pm$ 0,99	97,4 $\pm$ 1,55	12,6 $\pm$ 8,23	2,6 $\pm$ 1,55	87,4 $\pm$ 8,23
KNN	73,24 $\pm$ 0,58	91,85 $\pm$ 0,64	18,09 $\pm$ 2,42	8,15 $\pm$ 0,64	81,91 $\pm$ 2,42
LDA	62,82 $\pm$ 4,83	70,84 $\pm$ 9,35	39 $\pm$ 9,92	29,16 $\pm$ 9,35	61 $\pm$ 9,92

Da mesma forma que as demais, na Tabela 5 são apresentados os valores em ordem decrescente. Dessa forma, o classificador SVM teve o melhor desempenho médio entre os três, teve um TNR e TPR médios equivalente a $97,40\% \pm 1,55\%$ e $12,60\% \pm 8,23\%$, respectivamente.

5.1. TAREFA 1

A primeira tarefa tinha por objetivo realizar todos os treinamentos, considerando todas as possibilidades de classificadores, canais e indivíduos propostos nessa dissertação. Os resultados médios para exatidão são apresentados na Tabela 6.

Tabela 6 - Média de exatidão para cada classificador; Acc = exatidão; DP = Desvio Padrão. (O Autor, 2023).

ML	Média de Acc (%)	DP de Acc (%)
SVM	76,25	4,54
KNN	73,80	6,15
LDA	73,05	9,00

A Tabela 6 apresenta os valores calculados para cada um dos algoritmos classificadores utilizados. Vale destacar que essa média foi calculada a partir dos treinamentos individuais dos 15 voluntários de 30 e 90 tarefas para todos os canais e

os grupos coletivos, com 15 e 105 voluntários, com 90 tarefas de exemplo para todos os canais. Os valores apresentados já estão em ordem decrescente, ou seja, o melhor classificador foi o SVM, seguido do kNN e LDA.

A Tabela 7 apresenta a seleção de canais que teve o melhor desempenho médio contemplando todas as classificações realizadas até o momento. Assim sendo, o de 16 canais com $75,04\% \pm 6,85\%$ obteve o melhor desempenho médio e valores variando entre 53,30% até 90,00%. Nessa classificação, a de um canal está em quarta posição, com um resultado médio de $74,74\% \pm 6,85\%$ com valores variando de 56,70% até 96,70%.

Tabela 7 - Classificação dos resultados médios por quantidade de canais para todos os classificadores. Mín. Acc = Menor valor de exatidão obtido; Máx. Acc = Maior valor de exatidão obtido. (O Autor, 2023)

Seleção de Canais	Média de Acc (%)	DP de Acc (%)	Mín. Acc (%)	Máx. Acc (%)
16 Canais	75,04	6,85	53,30	90,00
9 Canais	74,92	6,79	56,70	93,30
19 Canais	74,86	6,37	57,70	90,00
1 Canal Cz	74,74	6,85	56,70	96,70
10 Canais	74,72	6,96	50,00	91,10
20 Canais	74,62	6,41	56,70	90,00
17 Canais	74,57	6,05	55,80	90,00
8 Canais	74,50	6,64	58,90	93,30
1 Canal FP1	74,41	8,90	46,70	96,70
6 Canais	74,14	7,08	53,30	90,00
5 Canais	73,68	6,37	56,70	93,30
14 Canais	73,31	7,88	43,30	93,30
4 Canais	73,27	6,85	54,40	90,00
Total Geral	74,37	6,95	43,30	96,70

Tendo em vista que a diferença entre os resultados obtidos é pequena, e considerando que o objetivo deste trabalho é estudar a redução de canais, será considerado o Canal Cz para realização das próximas tarefas. A Tabela 8 também apresenta de forma geral o desempenho de cada seleção de canais quando considerando apenas o SVM como classificador.

Tabela 8 - Classificação dos resultados médios por quantidade de canais para o classificador SVM. Mín. Acc = Menor valor de exatidão obtido; Máx. Acc = Maior valor de exatidão obtido. (O Autor, 2023)

Seleção de Canais	Média de Acc (%)	DP de Acc (%)	Mín. Acc (%)	Máx. Acc (%)
1 Canal Cz	77,44	4,92	70,00	91,10
1 Canal FP1	77,04	7,05	66,70	96,70
10 Canais	76,79	4,76	70,00	90,00
8 Canais	76,68	5,31	66,70	93,30
9 Canais	76,61	4,52	72,20	90,00
19 Canais	76,12	4,28	70,00	90,00
20 Canais	76,11	3,56	73,30	87,80
6 Canais	75,92	4,81	70,00	90,00
5 Canais	75,89	4,23	70,00	86,70
17 Canais	75,76	3,71	66,70	87,80
16 Canais	75,74	3,78	70,00	88,90
4 Canais	75,72	4,25	66,70	86,70
14 Canais	75,39	2,86	71,10	83,30
Total Geral	76,25	4,54	66,70	96,70

A Tabela 8 apresenta os resultados respectivos à quantidade de canais apenas realizados pelo classificador SVM e ela também está apresentada de forma decrescente. A quantidade de canais com a melhor performance é a de apenas um canal, sendo o melhor o Cz, seguido do Fp1. A tabela também apresenta os valores máximos e mínimos de exatidão obtidos durante os treinamentos.

Os que tiveram os mais baixos desempenhos médios foram a opção com 16 canais, com $75,74\% \pm 3,78\%$, seguido da opção com 4 canais, com $75,72\% \pm 4,25\%$, e por fim, com 14 canais, com $75,39\% \pm 2,86\%$.

A Tabela 8, portanto, reforça a escolha do canal Cz para realização das demais tarefas. Assim sendo, a seleção de configuração de canal visando a redução de canais e o melhor classificador a partir das médias de todos os treinamentos é 1 Canal Cz com SVM.

5.2. TAREFA 2

A tarefa 2 consistiu em, a partir do melhor classificador e configuração de canais, escolher o indivíduo com maior exatidão. Os resultados por indivíduo estão apresentados abaixo:

Tabela 9 - Valores máximos de exatidão obtidos pelo treinamento de um canal (Cz) para cada indivíduo para o algoritmo SVM (O Autor, 2023)

Sujeito x Tarefas	Máx. de Acc (%)
S015	91,10
30	73,30
90	91,10
S008	87,80
30	83,30
90	87,80
S004	85,60
30	83,30
90	85,60
S012	81,10
30	76,70
90	81,10
S003	81,10
30	73,30
90	81,10
S006	81,10
30	76,70
90	81,10
S010	80,00
30	76,70
90	80,00
S009	80,00
30	76,70
90	80,00
S001	77,80
30	70,00
90	77,80
S011	76,70
30	73,30
90	76,70
S013	76,70
30	76,70
90	76,70
S005	76,70
30	76,70
90	73,30
S007	74,40
30	73,30
90	74,40
S002	73,30
30	70,00
90	73,30
S014	73,30
30	73,30
90	73,30
Maior valor	91,10

A Tabela 9 mostra por ordem decrescente os valores máximos de exatidão encontrados, obtidos a partir dos treinamentos realizados para cada sujeito, sendo os iniciados com “S” seguidos de um número significando o treinamento individual de cada voluntário, correspondente aos treinamentos realizados com 30 exemplos de testes (tarefas) ou com 90.

Nessa análise é possível perceber que o sujeito com o melhor desempenho foi o sujeito S015, com a maior exatidão de 91,10% para o treinamento realizado com 90 exemplos de testes. Sendo assim, definiu-se a configuração de SVM com 1 Canal Cz com os dados coletados do sujeito S015 para demais treinamentos realizados durante as próximas tarefas.

5.3. TAREFA 3

A tarefa 3 teve por objetivo o teste do treinamento realizado com o arranjo de melhor classificador e canal escolhidos durante as tarefas 1 e 2. O exercício proposto é o teste das 60 demais tarefas para o treinamento realizado com 30 exemplos de testes. No treinamento de 30 tarefas do S015 o classificador apresentou uma exatidão de 73,30% com TNR de 100%, FPR de 0%, TPR de 0% e FNR de 100%.

O resultado obtido pelos testes é equivalente a 88,30% de exatidão para reconhecer 60 novas tarefas. Os resultados de TNR, FPR, TPR e FNR foram, respectivamente, 100%, 0%, 53,30% e 46,70%.

5.4. TAREFA 4

A tarefa 4 consistiu em, a partir da melhor configuração encontrada, treinar uma máquina com 60 exemplos de tarefas e testar as outras 30 restantes. Assim sendo, o resultado obtido foi de 80% de exatidão com TNR de 100%, FPR de 0%, TPR de 20% e FNR de 80% para o treinamento.

Ao testar esta máquina com 30 novos exemplos de tarefas o resultado da exatidão foi de 86,70%. Os resultados de TNR, FPR, TPR e FNR foram 100%, 0%, 50% e 50%, respectivamente.

5.5. TAREFA 5

Para a tarefa 5, foi realizada a redução de quantidade de exemplos de zeros nos exemplos de tarefas (T0 e T1) para o melhor indivíduo, na melhor configuração de canais selecionado na tarefa 1 e para todos os classificadores (SVM, LDA e kNN). Primeiramente, os resultados do treinamento inicial, considerando 1 Canal Cz para o sujeito S015 com 46 exemplos de tarefas, estão apresentados na Tabela 10.

Tabela 10 - Treinamento realizado com o sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 46 tarefas (O Autor, 2023).

ML	Acc (%)	TNR (%)	FPR (%)	TPR (%)	FNR (%)
SVM	84,80	91,30	8,70	78,30	21,70
LDA	78,30	78,30	21,70	78,30	21,70
KNN	78,30	78,30	21,70	78,30	21,70

A Tabela 10 apresenta os valores médios de exatidão e taxas de TNR, FPR, TPR e FNR, do treinamento inicial. Os valores de exatidão média para cada classificador foram 84,80%, 78,30% e 78,30% para SVM, LDA e kNN, respectivamente.

Tabela 11 - Treinamento realizado com o sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 90 tarefas (O Autor, 2023).

ML	Acc (%)	TNR (%)	FPR (%)	TPR (%)	FNR (%)
SVM	91,10	100,00	0,00	65,20	34,80
KNN	90,00	97,00	3,00	69,60	30,40
LDA	74,40	86,60	13,40	39,10	60,90

Com o objetivo de comparação, a Tabela 11 apresenta os valores médios de exatidão e taxas de TNR, FPR, TPR e FNR, do treinamento realizado com 90 tarefas, para o S015 com 1 Canal Cz. Os valores de exatidão média para cada classificador foram 91,10%, para SVM, 90,00% para LDA e 74,40% para kNN.

Tabela 12 – Teste da base de dados do sujeito S015 executando 90 tarefas realizado com o treinamento realizado para sujeito S015 com a configuração de canais 1 Canal Cz para todos os algoritmos classificadores com 46 tarefas (O Autor, 2023)

ML	Acc (%)	TNR (%)	FPR (%)	TPR (%)	FNR (%)
SVM	97,80	98,50	1,50	95,70	4,30
LDA	95,60	94,00	6,00	100,00	0,00
KNN	81,10	80,60	19,40	82,60	17,40

A Tabela 12 mostra os resultados obtidos quando o teste para classificar todas as 90 tarefas (na máquina treinada apenas com 46 exemplos) foi realizado. Ela mostra os valores de exatidão média para cada classificador, sendo eles, 97,80%, 95,60% e 81,10% para SVM, LDA e kNN, respectivamente.

6. DISCUSSÃO

O software MATLAB é uma ferramenta muito poderosa de análise de matrizes que vem se atualizando muito nos últimos anos com relação a aprendizado de máquinas e possibilitou que toda a análise de dados, classificação e testes fossem realizadas. A revisão sistemática realizada por Huamanchahua *et al.* (2021) informa que, para tecnologias assistivas baseadas em robôs paralelos, o software MATLAB é o mais utilizado devido à facilidade de processamento que o software dispõe. Após a análise dos dados coletados (programado internamente ao MATLAB), foi possível realizar os treinamentos e testes, sendo estes discutidos abaixo.

A partir da Tarefa 1, e assim como visto na revisão da literatura, o classificador que obteve o melhor desempenho acerca de todos os testes realizados foi o SVM, e isso foi comprovado como verdadeiro durante a realização desta pesquisa, tendo em vista que após todos os treinamentos aplicados na Tarefa 1, ele permaneceu com a melhor média geral.

Com relação aos canais, todos tiveram médias de exatidão bem parecidas e com um bom desempenho. Assim, visando selecionar o menor número de canais o suficiente para identificar movimento de membros inferiores, identificou-se que a seleção deveria ser composta por 1 canal, com a configuração de Cz.

Além de esta configuração apresentar-se com uma melhor média de exatidão quando comparado ao Fp1, Cz foi o canal mais utilizado quando analisado em todos os artigos da revisão, pois ele se localiza no centro do córtex motor primário que, conforme PURVES *et al.* (2008), tem a função de iniciar e executar movimento muscular voluntário. Essa região é responsável por iniciar todos os movimentos musculares, incluindo o de membros inferiores. Portanto, trata-se da principal região do cérebro que apresentará características no sinal capazes de distinguir a ação ou não.

Apesar disso, entende-se que é possível que variações ocorram de um indivíduo para outro, devido a fatores como sexo e dominância dos membros, já apontados em estudos para predição de movimentos de membro superior (CATRAMBONE *et al.*, 2019). Além disso, condições neurológicas já comprovaram afetar as ondas do sinal EEG para movimentos de membro superior parético

(BARTUR; PRATT; SOROKER, 2019). Portanto, deve-se considerar que a identificação do canal Cz em membro inferior também pode não ser representativo de modo universal.

Tendo sido feitas as considerações acerca dos fatores limitadores, pode-se destacar que, quando se observa os valores de exatidão obtidos com o melhor classificador (SVM), o canal que obteve a melhor média de exatidão foi justamente o Cz. Tal fato permite realizar o direcionamento do estudo a partir deste canal.

Três autores aplicaram a configuração de um canal apenas, sendo eles Jeong *et al.* (2020), Rodríguez-Ugarte *et al.* (2017), ambos estudando a configuração de 1 canal Cz, e Kuramoto *et al.* (2012), estudando a configuração de 1 Canal Fp1. Esses autores realizaram treinamento com diferentes algoritmos, incluindo o SVM, com a exceção do estudo feito por Jeong *et al.* (2020), que usou RLDA (Regularized Linear Discriminant Analysis) obtendo uma exatidão de 81%; porém, no presente estudo, por não se ter testado esse modelo de classificador, não houve como realizar a comparação. Os demais estudos utilizaram SVM.

Rodríguez-Ugarte *et al.* (2017) realizaram treinamentos de sinais EEG de membros inferiores e obtiveram resultados de até $84,60\% \pm 13,16\%$. Dessa forma, é possível confirmar que realmente é possível utilizar o Cz e obter bons valores de exatidão.

Por outro lado, Kuramoto *et al.* (2012) utilizaram o SVM para 1 Canal (Fp1). Os autores classificaram o sinal EEG de movimento de pés, obtendo uma exatidão de $89,80\% \pm 20\%$, que em comparação ao obtido neste estudo (máxima de $96,70\%$ para um voluntário específico), demonstra que, realmente, essa configuração é tão expressiva quanto utilizar o 1 Canal Cz em combinação com o SVM. Porém, quando comparado à média geral de $74,74\% \pm 6,85\%$ (4º melhor) para o Cz com $74,41\% \pm 8,90\%$ (9º melhor) para Fp1, percebe-se que o Fp1 está em uma colocação mais baixa.

Além disso, é importante destacar que estudos como o de Hortal *et al.* (2016) afirmam que, a partir de um estudo prévio realizado (HORTAL *et al.*, 2015), foi necessário aprimorar a metodologia desenvolvida e reduzir o tempo de preparação do experimento. Assim, a redução de canais (de 32 para 9) atenderia essa necessidade.

Além disso, percebeu-se uma maior aceitação dos usuários, tendo em vista o aumento do conforto durante o experimento proposto. Este fato reforça o benefício de se reduzir o número de canais a um número aceitável (relação redução x performance).

Ainda, Hortal *et al.* (2016) reforçam que essa redução de canais compreendeu a seleção somente na região do córtex motor primário, tendo em vista que aplicar em outras regiões poderia comprometer o sinal com artefatos que não são relevantes, além de pelo formato do capacete de eletrodos usado, o contato entre eletrodos e a cabeça não seria o ideal.

A Tarefa 3 foi proposta com o objetivo de verificar o quanto a máquina treinada consegue prever novos movimentos. Então, foram testados 60 novos exemplos de tarefas diferentes dos 30 exemplos de tarefas apresentadas para treinamento. Pode-se verificar que é possível prever novos movimentos com uma performance de 88,30%.

A Tarefa 4 foi similar à Tarefa 3; porém, a diferença é que foi disponibilizada uma maior quantidade de exemplos de teste em relação à quantidade de novos exemplos testados. A nova proporção realizada, nesta pesquisa, foi de 60 exemplos de execução para treinamento e 30 exemplos para teste da máquina. Assim, constatou-se que a máquina teve uma melhora na exatidão durante o treinamento de 60 tarefas (80% de exatidão) quando em comparação às 30 treinadas na tarefa anterior (73,30% de exatidão), totalizando um aumento de 6,70% de exatidão.

Durante o teste, foi possível perceber que o aumento da exatidão não foi obtido tendo-se em vista que a exatidão ao classificar novas amostras foi de 86,70%, valor menor ao obtido no teste da tarefa anterior (88,30%). Isso pode ser devido a uma possível fadiga no músculo alvo nas últimas 30 tarefas, alterando assim, levemente o sinal coletado e interferindo na classificação.

Estudos como Nugent *et al.* (2023) apontam que a fadiga muscular localizada é definida pela redução da força muscular induzida pela execução de exercícios, além disso, Deshpande e Hu (2011) realizaram um estudo que, por meio de ressonância magnética, foi possível verificar que existe uma redução da não linearidade derivada da diminuição da excitabilidade cortical dos neurônios. Sendo assim, há alterações que podem estar presentes no sinal EEG.

O objetivo principal subjacente à Tarefa 2 foi determinar a configuração a ser aplicada na metodologia das Tarefas 3 e 4 subsequentes que, por sua vez, aspiram em verificar os impactos do aumento de exemplos de coleta na classificação. Nenhum estudo seguindo essa metodologia foi encontrado na literatura. Porém seguindo a ideia apresentada pela Tarefa 4 de que a fadiga muscular atrapalha o sinal coletado, alguns autores acharam maneiras de fornecer novos exemplos de teste sem realizar a coleta. Um exemplo disso é o trabalho realizado por Rasheed *et al.* (2021).

Rasheed *et al.* (2021) objetivaram o aumento da base de dados de maneira sintética. A motivação deles era voltada à falta de dados EEG de boa qualidade para predição de crises epiléticas, que costumam ser escassas e difíceis de coletar leituras. Esses autores realizaram quatro experimentos aplicando o classificador SVM na base de dados de EEG: dois considerando treino da base de dados real testando ou a base real ou a base sintética aumentada, e dois considerando o treino da base de dados sintética aumentada testando ou a base real ou a base sintética aumentada. Ou seja, TRTR equivale a treino e teste da base de dados real, TSTR treino na base sintética e teste na base real, TRTS treino na base real e teste na base sintética e TSTS treino e teste na base sintética. Além disso, Rasheed *et al.* (2021) afirmam que os dados de treinamento são exemplos de realização de tarefas diferentes das apresentadas para testes, ou seja, um estudo muito similar ao desenvolvido nas Tarefas 3 e 4 dessa dissertação.

Os resultados de exatidão média obtidos por Rasheed *et al.* (2021) correspondem a: TRTR com 91,24% de exatidão média, TSTR com 90,03% de exatidão média, TRTS com 91,17% de exatidão média e TSTS com 92,02% de exatidão média. Entende-se que pelo fato de os autores terem treinado e testado a base de dados real, em períodos diferentes de execução, foi possível prever novos movimentos com uma alta exatidão, comprovando que o SVM consegue identificar o sinal solicitado com facilidade. Quando eles aplicaram a base de dados aumentada, o classificador atinge uma exatidão de 91,17%, muito similar ao valor encontrado para a base de dados original, mas com mais exemplos de momentos de crise, suportando novamente a conclusão obtida pelas Tarefas 3 e 4, que também obtiveram bons valores de exatidão para previsão de novos movimentos.

Analisando o valor obtido pelo treino e teste do modelo sintético aumentado de dados, equivalente a 92,02% de exatidão média, entende-se que com mais dados o desempenho do classificador melhora o modelo preditivo desenvolvido.

Como um dos objetivos da dissertação é identificar quando um indivíduo realiza uma execução motora em membros inferiores, ao fazer a análise dos dados, foram identificados mais exemplos de não execução ($T0$ e $T1 = 0$, 74,44% das tarefas em uma execução) do que de execuções ($T2 = 1$). Então, para verificar como se comportam as taxas de TPR e TNR, foram igualados os exemplos de tarefas em 50%/50% (23 exemplos de 0 e 1 cada, totalizando 46 exemplos de tarefas) na Tarefa 5.

Quando as máquinas treinadas com apenas 46 exemplos, são comparadas com a máquina treinada com os 90 exemplos de teste, é visível que não houve ganhos em relação a exatidão. Esses dados também reforçam as conclusões obtidas por Rasheed *et al.* (2021) de que com mais dados o desempenho do classificador melhora o modelo preditivo desenvolvido.

Analisando o TPR e TNR, que são os valores reais de classificação para realmente identificar ou não a execução de movimentos, respectivamente, por mais que as taxas gerais não foram melhores na classificação das tarefas para as máquinas treinadas com 46 tarefas, pode-se verificar que, em relação ao TPR (é a taxa que identifica a execução de movimentos), foram maiores em todos os testes. Ou seja, aumentando as proporções dos exemplos de execuções de movimentos ($T2$) em relação ao de não execução dos movimentos ($T0$ e $T1$), as máquinas foram mais precisas ao identificar a real execução de movimentos. Para essa análise não foram encontrados autores com trabalhos semelhantes.

É importante destacar que durante a realização deste trabalho, algumas limitações foram encontradas. Uma delas é a limitação apresentada pela base de dados disponibilizada, que apesar de ser uma base com muita informação possível de se trabalhar, ela não apresenta definições da demografia da população participante no estudo. Entretanto, com as informações disponibilizadas, pode-se assumir que se trata de pessoas hípidas, tendo em vista que precisaram executar funções motoras de pés e mãos com olhos abertos ou fechados.

Além disso, a base de dados informa que o sistema de BCI utilizado segue o sistema apresentado por Schalk *et al.* (2004), que define o sistema BCI2000. Em seu trabalho, os autores também não informam especificamente se foi aplicado algum tipo de filtro ou extrator de características, apenas apresentam recomendações de utilização. De toda forma, pode-se assumir que foi aplicado, devido a boa qualidade do sinal disponibilizado.

O hardware utilizado na realização deste estudo também se apresentou como uma limitação, tendo em vista que conforme o aumento da quantidade de canais/características, o uso de CPU e memória chegaram máximo, podendo assim, ser a causa do prolongamento do tempo de treinamento das máquinas.

7. CONCLUSÕES

Este estudo investigou várias combinações de treinamentos com algoritmos de aprendizado de máquina, com vistas a se identificar as combinações com o menor número de canais e o melhor desempenho na classificação de sinais de EEG em tarefas com membros inferiores de indivíduos.

A literatura consultada permitiu sintetizar o índice de desempenho de vários algoritmos utilizados para classificação de sinais EEG sobre BCIs para movimentos de membros inferiores. Para a população de indivíduos da base utilizada nesta pesquisa, o algoritmo SVM foi o classificador com maior exatidão média de predição, chegando a uma exatidão de até 91,10%. Além disso, com o uso de apenas 1 canal (Cz) é possível reconhecer movimentos em membros inferiores, com exatidão de até 96,70%.

Além disso, pode-se destacar a conclusão de que quanto maior for quantidade de exemplos para treinamento, melhor será a predição da máquina. Ainda, quanto mais exemplos da classe que se deseja identificar em relação à que não se deseja, maior será sua assertividade em reconhecê-las.

Esses fatos podem auxiliar a direcionar trabalhos futuros a decidir, por exemplo, qual a quantidade ideal de número de coletas que podem ser realizadas em um estudo, que sejam suficientes para a melhor predição da máquina, sem apresentar interferência da fadiga.

8. CONTRIBUIÇÕES

A realização deste trabalho permitiu definir as seguintes contribuições:

- Contribuição científica:
 - Determinação dentre os três algoritmos de classificação mais recorrentes que tiveram o melhor desempenho, em termos de exatidão, apresentados na literatura (SVM, LDA e kNN), qual teve a melhor performance para classificação de sinais EEG de movimento de membros inferiores. Sendo assim, a partir deste, não há necessidade de realizar novos estudos aplicando os demais;
 - Determinação de que a partir de apenas um canal (Cz), é suficiente para detecção de movimentos de membros inferiores a partir de sinais EEG e tem desempenho médio equivalente a utilizar maiores quantidades de canais;
 - Comprovação de que quanto mais exemplos de testes utilizados na hora do treinamento, maior será a exatidão da máquina ao reconhecer movimentos de membros inferiores;
 - Comprovação de que quanto mais exemplos de teste para uma tarefa a ser identificada em relação as tarefas a não serem identificadas, maior será a exatidão da máquina ao reconhecer as tarefas a serem identificadas.

9. TRABALHOS FUTUROS

São sugestões de trabalhos futuros:

- Realizar, por meio de coleta de sinais EEG, a análise semelhante à deste estudo em sinais de EEG de pessoas com diferentes condições neurológicas;
- Realizar, por meio de coleta de sinais EEG, a análise semelhante à deste estudo em sinais de EEG de voluntários com o controle da dominância do membro;
- Realizar, por meio de coleta de sinais EEG, a análise semelhante à deste estudo em sinais de EEG obtidos durante a execução de diferentes tarefas biomecânicas e/ou com membros superiores;
- Investigar métodos para melhorar a exatidão da configuração proposta a fim de aumentar a identificação de movimentos em membros inferiores;
- Validar se a aplicação de dados de sinais EEG com imagética motora, nas configurações determinadas como as de melhor desempenho durante a realização deste trabalho, trazem resultados semelhantes;
- Incorporar um sistema de BCI acoplável a uma órtese robótica de membro inferior, híbrida ou não, que considere os canais e utilize os algoritmos de classificação para aplicação *online*, com parâmetros identificados como de melhor desempenho durante a realização deste trabalho;
- Investigar uma possível quantidade ideal de coletas a ser realizada por um indivíduo, determinando assim uma possível quantidade suficiente para conseguir identificar o movimento com uma alta precisão;
- Realizar o mesmo estudo utilizando os métodos XBoost e Random Forest, tendo em vista sua ascensão como metodologias aplicadas na área, visando a comparação entre o emprego dos algoritmos.

REFERÊNCIAS BIBLIOGRÁFICAS

- BAI, Ou et al. A wireless, smart EEG system for volitional control of lower-limb prosthesis. In: TENCON 2015-2015 IEEE Region 10 Conference. IEEE, 2015. p. 1-6.
- BARTUR, Gadi; PRATT, Hillel; SOROKER, Nachum. Changes in mu and beta amplitude of the EEG during upper limb movement correlate with motor impairment and structural damage in subacute stroke. *Clinical Neurophysiology*, v. 130, n. 9, p. 1644-1651, 2019.
- BERRAR, Daniel. Cross-Validation. 2019.
- BERSCH, Rita de Cássia Reckziegel. Design de um serviço de tecnologia assistiva em escolas públicas. 2009.
- BIRBAUMER, Niels; COHEN, Leonardo G. Brain–computer interfaces: communication and restoration of movement in paralysis. *The Journal of physiology*, v. 579, n. 3, p. 621-636, 2007.
- BLANKERTZ, Benjamin et al. The BCI competition III: Validating alternative approaches to actual BCI problems. *IEEE transactions on neural systems and rehabilitation engineering*, v. 14, n. 2, p. 153-159, 2006.
- BONARDI, Gislaine; SOUZA, Valdemarina BA; MORAES, J. F. D. Incapacidade funcional e idosos: um desafio para os profissionais. *Scientia Medica*, v. 17, n. 3, p. 138-144, 2007.
- BRUNNER, P. et al. Current trends in hardware and software for brain–computer interfaces (BCIs). *Journal of neural engineering*, v. 8, n. 2, p. 025001, 2011.
- CATRAMBONE, Vincenzo et al. Predicting object-mediated gestures from brain activity: an EEG study on gender differences. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, v. 27, n. 3, p. 411-418, 2019.
- CHAN, Angela T. et al. An overview of brain computer interfaces. In: *Proc. 30th Int. Conf. on Computers and Their Applications*. 2015.
- CORTES, Corinna; VAPNIK, Vladimir. Support-vector networks. *Machine learning*, v. 20, p. 273-297, 1995.

CRISTIANINI, Nello et al. An introduction to support vector machines and other kernel-based learning methods. Cambridge university press, 2000.

DELISLE-RODRIGUEZ, Denis et al. System based on subject-specific bands to recognize pedaling motor imagery: towards a BCI for lower-limb rehabilitation. Journal of neural engineering, v. 16, n. 5, p. 056005, 2019.

DESHPANDE, Gopikrishna; HU, Xiaoping. Fatigue-induced changes in brain nonlinearity inferred by nonparametric and differential equation models of fMRI. In: 2011 IEEE International Conference on Industrial Technology. IEEE, 2011. p. 336-341.

DIAS, Raphael Mendes Ritti; GURJÃO, André Luiz Demantova; MARUCCI, Maria de Fátima Nunes. Benefícios do treinamento com pesos para aptidão física de idosos. Acta fisiátrica, v. 13, n. 2, p. 90-95, 2006.

DOS SANTOS, Edson G.; PAZ, Daniel R. Exojoelho. PROJETO FINAL EM ENGENHARIA DA COMPUTAÇÃO, 2018.

GHAHRAMANI, Zoubin. Unsupervised learning. In: Summer School on Machine Learning. Springer, Berlin, Heidelberg, 2003. p. 72-112.

GOLDBERGER, Ary L. et al. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. circulation, v. 101, n. 23, p. e215-e220, 2000.

HAYKIN, Simon. Redes neurais: princípios e prática. Bookman Editora, 2007.

HOLLAND, John H. et al. What is a learning classifier system?. In: International Workshop on Learning Classifier Systems. Springer, Berlin, Heidelberg, 1999. p. 3-32.

HORTAL, Enrique et al. EEG-based detection of starting and stopping during gait cycle. International journal of neural systems, v. 26, n. 07, p. 1650029, 2016.

HORTAL, Enrique et al. Using EEG signals to detect the intention of walking initiation and stop. In: Artificial Computation in Biology and Medicine: International Work-Conference on the Interplay Between Natural and Artificial Computation, IWINAC 2015, Elche, Spain, June 1-5, 2015, Proceedings, Part I 6. Springer International Publishing, 2015. p. 278-287.

HUAMANCHAHA, Deyby et al. Parallel Robots in Rehabilitation and Assistance: A Systematic Review. In: 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON). IEEE, 2021. p. 0692-0698.

JAIN, Anil K.; DUIN, Robert P.. W.. ; MAO, Jianchang. Statistical pattern recognition: A review. IEEE Transactions on pattern analysis and machine intelligence, v. 22, n. 1, p. 4-37, 2000.

JEONG, Ji-Hoon et al. Decoding movement-related cortical potentials based on subject-dependent and section-wise spectral filtering. IEEE Transactions on Neural Systems and Rehabilitation Engineering, v. 28, n. 3, p. 687-698, 2020.

KRAMER, Oliver. K-nearest neighbors. Dimensionality reduction with unsupervised nearest neighbors, 2013.

KUBLER, A. et al. BCI meeting 2005-workshop on clinical issues and applications. IEEE Transactions on neural systems and rehabilitation engineering, v. 14, n. 2, p. 131-134, 2006.

KURAMOTO, Nobuhisa et al. Trigger pattern detection method for assisting in ambulation rehabilitation based on EEG analysis. In: 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication. IEEE, 2012. p. 646-652.

LEE, Sangkyu; EL NAQA, Issam. Machine Learning in Radiation Oncology. Springer, p 3-11, 2015.

LI, Hanzhe et al. RP-based Voluntary Movement Intention Detection of Lower limb using CNN. In: 2020 10th Institute of Electrical and Electronics Engineers International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, 2020. p. 349-353.

MALIK, Anjum Naeem; IQBAL, Javaid; TIWANA, Mohsin I. EEG signals classification and determination of optimal feature-classifier combination for predicting the movement intent of lower limb. In: 2016 2nd International Conference on Robotics and Artificial Intelligence (ICRAI). IEEE, 2016. p. 45-49.

MAMMONE, Alessia; TURCHI, Marco; CRISTIANINI, Nello. Support vector machines. Wiley Interdisciplinary Reviews: Computational Statistics, v. 1, n. 3, p. 283-289, 2009.

MASKELIUNAS, Rytis et al. Consumer-grade EEG devices: are they usable for control tasks?. PeerJ, v. 4, p. e1746, 2016.

MATLAB & Simulink. Disponível em: <<https://www.mathworks.com/products/matlab.html>>. Acesso em: 09/06/2022

MDPI. Special "ssue "Advances in Artificial Intelligence and Machine Learning for B"/BMI". Disponível em: <https://www.mdpi.com/journal/bioengineering/special_issues/AI_BCI>. Acesso em: 12/02/2021

MOHRI, Mehryar; ROSTAMIZADEH, Afshin; TALWALKAR, Ameet. Foundations of machine learning. MIT press, 2018

NUGENT, Nicole S. et al. Investigating the effect of fatigue on muscle microvasculature blood flow during intermittent isometric contraction. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020. p. 3220-3223.

ONU – Nações Unidas. 1 bilhão de pessoas com deficiência entre as mais impactadas pela pandemia. 03 de Dezembro de 2021. Disponível em: <<https://news.un.org/pt/story/2021/12/1772482>>. Acesso em: 01/06/2022

PAGE, Matthew J. et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. International journal of surgery, v. 88, p. 105906, 2021.

PANOULAS, Konstantinos J.; HADJILEONTIADIS, Leontios J.; PANAS, Stavros M. Brain-computer interface (BCI): types, processing perspectives and applications. In: Multimedia services in intelligent environments. Springer, Berlin, Heidelberg, 2010. p. 299-321.

PLANELLES, Daniel et al. First steps in the development of an EEG-based system to detect intention of gait initiation. In: 2014 IEEE International Systems Conference Proceedings. IEEE, 2014. p. 167-171.

Política Nacional de Saúde da Pessoa com Deficiência. Disponível em: <<https://bvsms.saude.gov.br/bvs/saudelegis/gm/2017/MatrizesConsolidacao/comum/37518.html#:~:text=As%20pessoas%20com%20defici%C3%Aancia%20motora,m%>>

C3%A1%2Dforma%C3%A7%C3%A3o%20cong%C3%AA Anita%20ou%20adquirida.>
Acesso em: 16/04/2023

PRATT, Jerry E. et al. The RoboKnee: an exoskeleton for enhancing strength and endurance during walking. In: IEEE International Conference on Robotics and Automation, 2004. Proceedings' ICRA'04. 2004. IEEE, 2004.

PURVES, Dale et al. Modification of brain circuits as a result of experience. Neuroscience. 4th ed. Sunderland, MA: Sinauer Associates Inc, p. 611-633, 2008.

RASHEED, Khansa et al. A generative model to synthesize eeg data for epileptic seizure prediction. IEEE Transactions on Neural Systems and Rehabilitation Engineering, v. 29, p. 2322-2332, 2021.

RODRÍGUEZ-UGARTE, M. et al. Preliminary study of pedaling motor imagery classification based on EEG signals. In: 2017 International Symposium on Wearable Robotics and Rehabilitation (WeRob). IEEE, 2017. p. 1-2.

RODRÍGUEZ-UGARTE, Marisol et al. Personalized offline and pseudo-online BCI models to detect pedaling intent. Frontiers in neuroinformatics, v. 11, p. 45, 2017.

SANEI, Saeid; CHAMBERS, Jonathon A. EEG signal processing. John Wiley & Sons, 2013.

SAZGAR, Mona et al. Overview of EEG, electrode placement, and montages. Absolute epilepsy and EEG rotation review: essentials for trainees, p. 117-125, 2019.

SCHALK, Gerwin et al. BCI2000: a general-purpose brain-computer interface (BCI) system. IEEE Transactions on biomedical engineering, v. 51, n. 6, p. 1034-1043, 2004.

SHASHOA, Nasar Aldian Ambark et al. Classification depend on linear discriminant analysis using desired outputs. In: 2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA). IEEE, 2016. p. 328-332.

SOMADDER, Rittika; SAHA, Dabasish Kumar. Frequency Domain CSP for Foot Motor Imagery Classification Using SVM for BCI Application. In: 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES). IEEE, 2021. p. 30-34.

SORIANO-SEGURA, Paula et al. Selection of Spatial, Temporal and Frequency Features to Detect Direction Changes during Gait. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2020. p. 3835-3838.

TEIXEIRA, Luzimar. Deficiência física. (2010). Disponível em: <<http://www.luzimarteixeira.com.br/wp-content/uploads/2010/05/definicao-e-classificacao-da-deficiencia-fisica.pdf>>. Acesso em: 16/04/2023

TRIANA-GUZMAN, Nayid et al. Decoding EEG rhythms offline and online during motor imagery for standing and sitting based on a brain-computer interface. *Frontiers in Neuroinformatics*, 2022.

WOLPAW, Jonathan R. et al. BCI meeting 2005-workshop on signals and recording methods. *IEEE Transactions on neural systems and rehabilitation engineering*, v. 14, n. 2, p. 138-141, 2006.

WOLPAW, Jonathan R. et al. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, v. 113, n. 6, p. 767-791, 2002.

XU, Ren et al. A closed-loop brain-computer interface triggering an active ankle-foot orthosis for inducing cortical neural plasticity. *IEEE Transactions on Biomedical Engineering*, v. 61, n. 7, p. 2092-2101, 2014.