



UNIVERSIDADE FEDERAL DO CEARÁ
CAMPUS QUIXADÁ
CURSO DE GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

JOÃO VICTOR AQUINO CORREIA

**UTILIZANDO APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DE
DIAGNÓSTICOS DA DOENÇA DE ALZHEIMER**

QUIXADÁ

2023

JOÃO VICTOR AQUINO CORREIA

UTILIZANDO APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DE DIAGNÓSTICOS
DA DOENÇA DE ALZHEIMER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação do CAMPUS QUIXADÁ da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Ciência da Computação.

Orientador: Prof. Dr. Regis Pires Magalhães.

QUIXADÁ

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

C848u Correia, João Victor Aquino.
Utilizando Aprendizado de Máquina para Predição de Diagnósticos da Doença de Alzheimer / João Victor Aquino Correia. – 2023.
62 f. : il. color.

Trabalho de Conclusão de Curso (graduação) – Universidade Federal do Ceará, , Quixadá, 2023.
Orientação: Prof. Dr. Regis Pires Magalhães.

1. Aprendizado de máquina. 2. Aprendizado Profundo. 3. Redes Neurais. 4. Doença de Alzheimer. 5. Demência. I. Título.

CDD

JOÃO VICTOR AQUINO CORREIA

UTILIZANDO APRENDIZADO DE MÁQUINA PARA PREDIÇÃO DE DIAGNÓSTICOS
DA DOENÇA DE ALZHEIMER

Trabalho de Conclusão de Curso apresentado ao Curso de Graduação em Ciência da Computação do CAMPUS QUIXADÁ da Universidade Federal do Ceará, como requisito parcial à obtenção do grau de bacharel em Ciência da Computação.

Aprovada em: 14 de Julho de 2023.

BANCA EXAMINADORA

Prof. Dr. Regis Pires Magalhães (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Criston Pereira de Souza
Universidade Federal do Ceará (UFC)

Profa. Dra. Lívia Almada Cruz
Universidade Federal do Ceará (UFC)

A Deus, meu guia.

AGRADECIMENTOS

Primeiramente, agradeço a Deus, pois sem Ele não teria sido possível completar esta etapa. Em segundo lugar, agradeço aos que sempre estiveram comigo, minha família. Agradeço a minha mãe, Maria Kilma, por acreditar e apoiar nas minhas decisões. Agradeço ao meu Pai, Silvino, homem trabalhador e esforçado.

Agradeço ao meu orientador, Prof. Dr. Regis Pires Magalhães, um exemplo de profissional e um professor de grande conhecimento. Agradeço por aceitar o convite e por toda ajuda prestada. Agradeço por não desistir de mim nos momentos mais difíceis.

Agradeço ao Narcelio de Oliveira, Romes Filho, Rayrisson Vinícius, Vito Neto, José Deyvidson e Carlos Ryan, amigos do ensino médio e companheiros de apartamento em Quixadá, não sei como suportei vocês por todo esse tempo, estou brincando, vocês são incríveis.

Agradeço a Alysson Araújo, João Almir, Diogo Sampaio, Álisson Vêras, Andreina Mendes, Tiago Rodrigues, Luis Fernando, Pedro Spinosa, amigos que conheci na faculdade, vocês são demais. Agradeço a outras pessoas que não foram citadas, a turma de CC 2019.1, os professores, em especial a Paulyne Matheus Jucá, excelente professora e uma pessoa de grande coração.

Preciso agradecer de forma especial para algumas pessoas. Elano Nunes, um grande amigo, companheiro de trabalhos acadêmicos, de piadas e de momentos memoráveis. Mariana Oliveira, uma amiga que esteve comigo e com o Elano desde o início. Eduarda Eça, uma magnífica amiga e ótima namorada, agradeço pelo seu amor e por todos os momentos que passamos juntos, você está no meu coração. Por fim, agradeço as pessoas que fazem parte da Universidade Federal do Ceará, os professores participantes da banca examinadora e outras pessoas que contribuíram para a realização deste momento.

RESUMO

A doença de Alzheimer é uma doença neurodegenerativa gradual e irreversível, de aparecimento insidioso, que causa perda de memória e uma variedade de problemas cognitivos. Atualmente, é o tipo de demência mais comum no mundo, representando cerca de 60% a 80% dos casos. Não há uma forma simples de detectar. O diagnóstico requer um exame médico completo. O diagnóstico precoce dá ao paciente e à família uma chance maior de retardar a progressão da doença e ter mais controle sobre os sintomas. Embora não haja cura ou uma maneira de interromper sua progressão e o tratamento ainda seja uma questão de pesquisa em aberto, existem opções farmacológicas e não farmacológicas que podem ajudar a tratar os sintomas. A detecção precoce por meio da ressonância magnética convencional facilitará intervenções/tratamentos eficazes em tempo hábil, que ampliariam a expectativa e a qualidade de vida dos pacientes. Aprendizado de máquina é uma área de pesquisa da ciência da computação que envolve a indução de hipóteses ou funções, a partir de experiências passadas, para resolver um determinado problema de natureza preditiva ou descritiva. Com o avanço do desenvolvimento da tecnologia de aprendizado de máquina, ele pode ser usado como um rápido método auxiliar de diagnóstico. O aprendizado profundo é uma subárea do aprendizado de máquina, pode ser pensado como uma forma automatizada de análise preditiva, que tem como ênfase sucessivas camadas com representações cada vez mais significativas. Este trabalho propõe auxiliar no diagnóstico preciso da doença de Alzheimer utilizando o aprendizado de máquina e o aprendizado profundo. Este trabalho tem como principal objetivo propor e avaliar uma solução para a detecção de diagnósticos da doença de Alzheimer em imagens de ressonância magnética, utilizando técnicas de aprendizagem de máquina e aprendizado profundo. Desta forma, foram implementados ao todo 15 modelos de aprendizado de máquina e 4 diferentes arquiteturas de redes neurais convolucionais.

Palavras-chave: Aprendizado de máquina; Aprendizado Profundo, Redes Neurais; Doença de Alzheimer.

ABSTRACT

Alzheimer's disease is a gradual and irreversible neurodegenerative disease with an insidious onset that causes memory loss and a variety of cognitive problems. Currently, it is the most common type of dementia in the world, representing about 60% to 80% of cases. There is no simple way to detect it. Diagnosis requires a complete medical examination. Early diagnosis gives the patient and family a better chance of slowing the progression of the disease and having more control over the symptoms. While there is no cure or way to stop its progression and treatment is still an open research question, there are pharmacological and non-pharmacological options that can help treat the symptoms. Early detection through conventional magnetic resonance imaging will facilitate timely and effective interventions/treatments, which would increase patients' life expectancy and quality of life. Machine learning is an area of computer science research that involves the induction of hypotheses or functions, from past experiences, to solve a given problem of a predictive or descriptive nature. With the advancement of machine learning technology development, it can be used as a quick auxiliary diagnostic method. Deep learning is a sub-area of machine learning, it can be thought of as an automated form of predictive analysis, which emphasizes successive layers with increasingly meaningful representations. This work proposes to assist in the accurate diagnosis of Alzheimer's disease using machine learning and deep learning. The main objective of this work is to propose and evaluate a solution for the detection of Alzheimer's disease diagnoses in magnetic resonance images, using machine learning and deep learning techniques. In this way, 15 machine learning models and 4 different convolutional neural network architectures were implemented.

Keywords: Machine learning; Deep Learning, Neural Networks; Alzheimer's disease.

LISTA DE ILUSTRAÇÕES

Figura 1 – Rede neural simples e rede neural profunda	30
Figura 2 – Fluxo de procedimentos metodológicos	39
Figura 3 – Imagem de ressonância magnética da classe demente muito leve	45
Figura 4 – Imagem de ressonância magnética da classe demente leve	45
Figura 5 – Imagem de ressonância magnética da classe demente moderado	46
Figura 6 – Imagem de ressonância magnética da classe não demente	46
Figura 7 – Arquitetura rede neural convulacional 01	49
Figura 8 – Arquitetura rede neural convulacional 02	50
Figura 9 – Arquitetura rede neural convulacional 03	57
Figura 10 – Arquitetura rede neural convulacional 04	58

LISTA DE TABELAS

Tabela 1 – Experimento 1, avaliação dos modelos com hiperparâmetros padrão	51
Tabela 2 – Experimento 2, avaliação dos modelos utilizando a validação cruzada	52
Tabela 3 – Experimento 3, avaliação dos modelos utilizando a SMOTE e com hiperparâ- metros padrão	53
Tabela 4 – Experimento 4, avaliação dos modelos utilizando a técnica <i>gridSearch</i> . . .	54
Tabela 5 – Hiperparâmetros do modelo árvore de decisão	54
Tabela 6 – Hiperparâmetros do modelo árvores extras	55
Tabela 7 – Hiperparâmetros do modelo floresta aleatória	55
Tabela 8 – Hiperparâmetros do modelo <i>adaBoost</i>	55
Tabela 9 – Experimento 5, avaliação dos modelos de aprendizado profundo	55

LISTA DE QUADROS

Quadro 1 – Quadro comparativo entre os trabalhos relacionados e o trabalho proposto .	37
Quadro 2 – Distribuição das classes nos conjuntos de dados	44

LISTA DE ABREVIATURAS E SIGLAS

ADNI	Iniciativa de neuroimagem da doença de Alzheimer
AM	Aprendizado de máquina
AP	Aprendizado profundo
CNN	<i>Convolutional neural network</i> / Rede neural convolucional
FN	Falsos negativos
FP	Falsos positivos
KNN	K-vizinhos mais próximos
MIRIAD	<i>Minimal interval resonance imaging in Alzheimer's disease</i>
OASIS	<i>Open access series of imaging studies</i>
RVA	<i>Recurrent visual attention</i> / Atenção visual recorrente
SGD	<i>Stochastic gradient descent</i> / Gradiente descendente estocástico
SMOTE	Técnica de sobreamostragem minoritária sintética
SVM	Máquina de vetores de suporte
VN	Verdadeiros negativos
VP	Verdadeiros positivos

SUMÁRIO

1	INTRODUÇÃO	14
1.1	Objetivos	15
<i>1.1.1</i>	<i>Objetivo Geral</i>	15
<i>1.1.2</i>	<i>Objetivos Específicos</i>	16
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Doença de Alzheimer	17
<i>2.1.1</i>	<i>Alterações Cerebrais</i>	18
<i>2.1.2</i>	<i>Demência</i>	18
<i>2.1.3</i>	<i>Sintomas</i>	19
<i>2.1.4</i>	<i>Fases</i>	19
<i>2.1.5</i>	<i>Diagnóstico</i>	21
<i>2.1.6</i>	<i>Tratamento</i>	21
2.2	Aprendizado de Máquina	22
<i>2.2.1</i>	<i>Aprendizado Supervisionado</i>	23
<i>2.2.2</i>	<i>Classificação</i>	23
<i>2.2.3</i>	<i>Árvore de Decisão</i>	23
<i>2.2.4</i>	<i>Árvores Extremamente Aleatórias</i>	24
<i>2.2.5</i>	<i>Floresta Aleatória</i>	24
<i>2.2.6</i>	<i>KNN</i>	24
<i>2.2.7</i>	<i>SGD</i>	25
<i>2.2.8</i>	<i>Regressão Logística</i>	25
<i>2.2.9</i>	<i>Naive Bayes</i>	26
<i>2.2.10</i>	<i>SVM</i>	26
<i>2.2.11</i>	<i>GradientBoosting</i>	26
<i>2.2.12</i>	<i>XGBoost</i>	27
<i>2.2.13</i>	<i>HistGradientBoosting</i>	27
<i>2.2.14</i>	<i>Análise Discriminante Linear</i>	27
<i>2.2.15</i>	<i>Ridge</i>	28
2.3	Balanceamento	28
<i>2.3.1</i>	<i>SMOTE</i>	28

2.4	Aprendizado Profundo	29
2.4.1	<i>Redes Neurais</i>	29
2.4.2	<i>Redes Neurais Convolucionais</i>	30
2.4.3	<i>Early Stopping</i>	31
2.5	Métricas de Avaliação	31
2.5.1	<i>Acurácia</i>	32
2.5.2	<i>Precisão</i>	32
2.5.3	<i>Revocação</i>	32
2.5.4	<i>F1-score</i>	32
3	TRABALHOS RELACIONADOS	34
3.1	<i>Model Based on Deep Feature Extraction for Diagnosis of Alzheimer’s Disease</i>	34
3.2	<i>Preclinical Stage Alzheimer’s Disease Detection Using Magnetic Resonance Image Scans</i>	35
3.3	<i>DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images</i>	36
3.4	<i>Understanding Convolutional Neural Network’s Behavior for Alzheimer’s Disease on MRI</i>	36
3.5	Análise Comparativa	37
4	PROCEDIMENTOS METODOLÓGICOS	39
4.1	Coleta do Conjunto de Dados	39
4.2	Pré-processamento e análise de Dados	39
4.3	Particionamento dos Dados em Conjunto de Treino, Teste e Validação	40
4.4	Seleção dos Modelos de Aprendizado Máquina e Aprendizado Profundo	40
4.5	Treinamento e Otimização dos Hiperparâmetros	40
4.6	Avaliação dos Modelos	41
5	EXPERIMENTOS E RESULTADOS	42
5.1	Coleta do Conjunto de Dados	42
5.2	Pré-processamento e Análise de Dados	43
5.2.1	<i>Carregamento dos Dados</i>	43
5.2.2	<i>Particionamento dos Dados em Conjunto de Treino, Teste e Validação</i>	43
5.2.3	<i>Normalização e Codificação de Variáveis Categóricas</i>	44

5.2.4	<i>Análise de Dados</i>	44
5.3	Seleção dos Modelos de Aprendizado de Máquina e Aprendizado Profundo	44
5.4	Treinamento e Otimização dos Hiperparâmetros	47
5.4.1	<i>Arquitetura das Redes Neurais</i>	47
5.5	Avaliação dos Modelos	48
5.5.1	<i>Cenários de Experimentação</i>	48
5.5.1.1	<i>Experimento 1</i>	49
5.5.1.2	<i>Experimento 2</i>	51
5.5.1.3	<i>Experimento 3</i>	52
5.5.1.4	<i>Experimento 4</i>	54
5.5.1.5	<i>Experimento 5</i>	54
5.6	Interpretação e Análise dos Resultados	56
6	CONCLUSÕES	59
	REFERÊNCIAS	60

1 INTRODUÇÃO

A doença de Alzheimer, identificada pela primeira vez pelo neuropatologista alemão Alois Alzheimer em 1907, é uma doença neurodegenerativa gradual e irreversível, de aparecimento insidioso, que causa perda de memória e uma variedade de problemas cognitivos (SMITH, 1999). É caracterizada por um declínio progressivo na memória, raciocínio, compreensão, capacidade de computação matemática, linguagem, capacidade de aprendizado e julgamento, o que acaba tornando a pessoa afetada incapaz de realizar tarefas diárias independentemente (BARROS *et al.*, 2009).

A progressão da doença ocasiona desde alterações cerebrais imperceptíveis para a pessoa afetada até alterações cerebrais que causam problemas de memória e, eventualmente, incapacidade física. Existem três fases: pré-clínica, que ocorre antes do aparecimento de qualquer sintoma; comprometimento cognitivo leve, quando o paciente começa a apresentar comprometimento cognitivo, mas não impede que os pacientes realizem todas as atividades diárias; demência de Alzheimer, também chamada de demência, que ocorre quando os sintomas são graves o suficiente para impedir que os pacientes realizem as atividades diárias. A fase de demência é subdividida em demência leve, moderada e grave (GAUGLER *et al.*, 2022).

Atualmente, a doença de Alzheimer é o tipo de demência mais comum no mundo, representando cerca de 60% a 80% dos casos (GAUGLER *et al.*, 2022). Estima-se que em todo o mundo, o Alzheimer afeta 35,6 milhões de pacientes. Calcula-se que até 2050 o número de pacientes que sofrem de Alzheimer e outras demências aumente para 115,4 milhões de casos (CHEN; ZHONG, 2013). No Brasil estima-se que existam 1,2 milhão de casos, a maioria deles ainda sem diagnóstico (ABRAZ, 2022).

Um em cada três idosos morre com doença de Alzheimer ou outra demência. Ela mata mais do que o câncer de mama e o câncer de próstata juntos. As mortes mais do que dobraram entre 2000 e 2019. Aos 70 anos, os idosos que vivem com a doença têm duas vezes mais chances de morrer antes dos 80 anos do que aqueles que não têm a doença (GAUGLER *et al.*, 2022). Pessoas com 65 anos ou mais sobrevivem em média de quatro a oito anos após o diagnóstico, mas algumas vivem até 20 anos (GAUGLER *et al.*, 2022).

Não há uma forma simples de detectar a doença. O diagnóstico requer um exame médico completo, o qual pode incluir: histórico médico familiar, exames de sangue, testes cognitivos e imagiologia cerebral podem ser usados para determinar a causa dos sintomas (ALZ, 2022). O diagnóstico precoce dá ao paciente e à família uma chance maior de retardar a

progressão da doença e ter mais controle sobre os sintomas (ABRAZ, 2022). Embora não haja cura ou uma maneira de interromper sua progressão e o tratamento ainda seja uma questão de pesquisa em aberto, existem opções farmacológicas e não farmacológicas que podem ajudar a tratar os sintomas. A detecção precoce por meio da ressonância magnética convencional facilitará intervenções e tratamentos eficazes em tempo hábil, que ampliariam a expectativa e a qualidade de vida dos pacientes (ALTAY *et al.*, 2021).

O Aprendizado de máquina (AM) é uma área de pesquisa da ciência da computação que envolve a indução de hipóteses ou funções, a partir de experiências passadas, para resolver um determinado problema de natureza preditiva ou descritiva (FACELI *et al.*, 2011). O AM tornou-se rapidamente o subcampo mais popular e bem-sucedido da inteligência artificial, uma tendência impulsionada pela disponibilidade de *hardware* mais rápido e conjuntos de dados maiores (CHOLLET, 2021). Com o avanço do desenvolvimento da tecnologia de aprendizado de máquina, as pessoas descobriram que ele pode ser usado como um rápido método auxiliar de diagnóstico (WU *et al.*, 2022).

O Aprendizado profundo (AP) é uma subárea do aprendizado de máquina, que emprega algoritmos para processar dados e imita o método de aprendizado que as pessoas usam para aprender tipos específicos de conhecimento, é um sistema baseado em neurônios (CHOLLET, 2021). O AP pode ser pensado como uma forma automatizada de análise preditiva, que tem como ênfase sucessivas camadas com representações cada vez mais significativas (CHOLLET, 2021). Ele é mais adequado para questões desafiadoras, como reconhecimento de imagem e voz ou processamento de linguagem natural, contanto que tenha dados suficientes, capacidade de processamento e paciência (GÉRON, 2019).

Portanto, este trabalho visa auxiliar no diagnóstico preciso da doença de Alzheimer, considerando o potencial uso do aprendizado de máquina e profundo para auxiliar na detecção de eventos relacionados à saúde, na coleta de informações e na redução de riscos por meio de diagnósticos.

1.1 Objetivos

1.1.1 Objetivo Geral

O objetivo principal deste trabalho é treinar e avaliar modelos de aprendizado de máquina e aprendizado profundo para predição de diagnósticos da doença de Alzheimer nas

classes demente muito leve, demente leve, demente moderado e não demente, usando imagens de ressonância magnética.

1.1.2 Objetivos Específicos

- Avaliar e comparar o desempenho dos modelos utilizando imagens na escala de cinza e na escala *RGB*
- Avaliar e comparar o desempenho dos modelos utilizando técnicas de balanceamento de dados e não utilizando
- Avaliar e comparar o desempenho dos modelos usando diversas métricas de classificação
- Comparar estratégias de aprendizado profundo com aprendizado de máquina para a classificação multiclasse das imagens de ressonâncias magnéticas

Os próximos capítulos estão organizados da seguinte maneira: o Capítulo 2 apresenta a fundamentação teórica e os conceitos que alicerçam as abordagens propostas neste trabalho. O Capítulo 3 trata dos trabalhos relacionados, são descritos e comparados os projetos e pesquisas que possuem aspectos similares aos especificados neste trabalho. o Capítulo 4 apresenta-se a metodologia com a descrição das atividades relacionadas ao desenvolvimento e validação dos modelos de classificação propostos. O Capítulo 5 apresenta os experimentos e resultados. E por fim, no Capítulo 6 é apresentada as conclusões.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo são apresentados os conceitos que fundamentam o desenvolvimento deste trabalho. Inicialmente, a Seção 2.1 apresenta a descrição da doença de Alzheimer e suas características neurobiológicas, assim como alguns critérios de diagnósticos e tratamentos. A Seção 2.2 traz os conceitos e características referentes ao aprendizado de máquina. Na Seção 2.4 são apresentados os conceitos de aprendizado profundo. Por fim, a Seção 2.5 apresenta as métricas adotadas para fins de avaliação dos algoritmos de classificação.

2.1 Doença de Alzheimer

A doença de Alzheimer é um tipo de doença cerebral, causada por danos às células nervosas (neurônios). Indivíduos com sintomas leves normalmente podem continuar executando suas atividades diárias, como, por exemplo: trabalhar, dirigir e participar de suas atividades favoritas com a ajuda ocasional de amigos e familiares. No entanto, ela é uma doença progressiva, o que significa que piora com o tempo. A rapidez com que progride e quais habilidades são afetadas variam de pessoa para pessoa. Os medicamentos podem ajudar temporariamente os neurônios a se comunicarem mutuamente e, dessa forma, ajudar os sintomas por períodos variados, mas não curam a doença (GAUGLER *et al.*, 2022).

A doença de Alzheimer é uma patologia incurável, quase todas as suas vítimas são idosos. A doença apresenta-se como demência, ou a perda de habilidades cognitivas, provocada pela morte de células cerebrais. Quando diagnosticado precocemente, é possível retardar a progressão e exercer maior controle sobre os sintomas, garantindo uma melhor qualidade de vida ao paciente e à família (ABRAZ, 2022).

Os primeiros neurônios danificados são aqueles nas regiões cerebrais responsáveis pela memória, linguagem e pensamento. Como resultado, os primeiros sintomas tendem a ser problemas de memória, linguagem e pensamento. Embora esses sintomas possam ser novos para a pessoa afetada, acredita-se que as alterações cerebrais que os levaram começaram há pelo menos vinte anos antes. Com o passar do tempo, mais neurônios são danificados e mais áreas do cérebro são afetadas. É necessária maior ajuda de familiares, amigos e cuidadores profissionais para atividades diárias como vestir-se e usar o banheiro, bem como para manter o indivíduo seguro. As pessoas podem experimentar mudanças de humor, comportamento ou personalidade como resultado dos desafios mentais e de memória que estão enfrentando. Uma área particular

de preocupação é a prevenção de quedas, que podem resultar em lesões cranianas, fraturas e hospitalização. Eventualmente, o dano causado estende-se a áreas do cérebro que permitem funções corporais básicas, como caminhar e segurar. Os indivíduos ficam angustiados e precisam de cuidados o dia inteiro. Em último caso, a doença é fatal. Segundo estudos, pessoas com 65 anos ou mais sobrevivem em média de quatro a oito anos após serem diagnosticadas com Alzheimer, enquanto outras chegam a viver por até vinte anos (GAUGLER *et al.*, 2022).

2.1.1 Alterações Cerebrais

Embora a causa da doença de Alzheimer seja desconhecida, algumas lesões cerebrais características dessa doença são conhecidas. As duas principais alterações que ocorrem são as placas senis decorrentes do depósito anormal da proteína beta-amilóide produzida e os neurofibrilares emaranhados causados pela hiperfosforilação da proteína tau. Outra mudança que tem sido notada é a redução do número de células nervosas e das conexões entre elas (sinapses), com redução progressiva do volume cerebral. Estudos recentes mostraram que essas alterações cerebrais já existiam antes do início dos sintomas degenerativos (ABRAZ, 2022).

A doença de Alzheimer é uma doença cerebral de progressão lenta que se desenvolve muitos anos antes do início dos sintomas. Ela representa entre 60% e 80% dos casos de demência, tornando-a a causa mais frequente. Estudos recentes de autópsia revelam que mais da metade das pessoas com doença exibem alterações cerebrais causadas pela própria doença de Alzheimer, bem como alterações causadas por uma ou mais causas adicionais de demência, como doença cerebrovascular ou doença de corpos de Lewy (GAUGLER *et al.*, 2022).

2.1.2 Demência

Demência é um termo geral para um conjunto específico de sintomas. A demência é caracterizada por dificuldades de memória, linguagem, habilidades de resolução de problemas e outras habilidades mentais. Existem várias causas para a demência. Essas causas refletem mudanças específicas no cérebro. A causa mais frequente de demência são as alterações cerebrais relacionadas à doença de Alzheimer (GAUGLER *et al.*, 2022).

2.1.3 Sintomas

Costumam ser sintomas precoce de Alzheimer: dificuldade em lembrar conversas, nomes ou eventos recentes, apatia e depressão. Os sintomas posteriores incluem comunicação prejudicada, desorientação, confusão, falta de julgamento, alterações comportamentais e, finalmente, dificuldade para falar, engolir e andar (GAUGLER *et al.*, 2022).

Os pacientes frequentemente experimentam perda de memória recente com a repetição das mesmas perguntas ou tópicos. Esquecimento de eventos, de compromissos ou local onde se guardava seus pertences. Dificuldade em reconhecer uma situação de risco, cuidar das próprias finanças e pertences pessoais, tomar decisões e planejar atividades mais complexas. Dificuldade de orientação espacial e temporal. Incapacidade de reconhecer rostos ou objetos familiares, possivelmente dificultando o reconhecimento de pessoas conhecidas (ABRAZ, 2022).

Além disso, os paciente tem dificuldade para manusear utensílios, vestir-se e realizar atividades que exijam autocuidado. Dificuldade em encontrar e/ou compreender palavras, cometendo erros ao falar e escrever. Mudanças no comportamento ou na personalidade incluem tornar-se agitado, apático, desinteressado, isolado, desinibido, inapropriado ou mesmo agressivo (ABRAZ, 2022).

Outros sintomas incluem percepções exageradas da realidade, como quadros paranoicos ao achar que está sendo roubado, perseguido ou enganado por alguém. Esquecer o que aconteceu ou o que ficou combinado pode contribuir para esse quadro. Podem ocorrer alucinações visuais ou auditivas, sendo mais frequentes da metade para o final do dia. Alterações no apetite que podem levar a comer demais ou, inversamente, podem ocorrer uma diminuição do apetite. Atividade noturna ou insônia com transição do dia para a noite (ABRAZ, 2022).

2.1.4 Fases

Existem três fases da doença de Alzheimer: pré-clínica, que ocorre antes do aparecimento de qualquer sintoma. Comprometimento cognitivo leve, quando o paciente começa a apresentar comprometimento cognitivo, mas não impede que os pacientes realizem todas as atividades diárias. E demência, também chamada de demência de Alzheimer, que ocorre quando os sintomas são graves o suficiente para impedir que os pacientes realizem as atividades da vida diária. A fase de demência é subdividida em demência leve, moderada e grave (GAUGLER *et al.*, 2022). Abaixo são descritos cada fase de acordo com Gaugler *et al.* (2022).

- **Pré-clínica:** nesta fase, as pessoas podem ver mudanças mensuráveis em seus cérebros que representam os primeiros sintomas da doença, mas ainda não experimentaram sintomas como perda de memória
- **Comprometimento cognitivo leve:** nesta fase, os pacientes já apresentam biomarcadores de alterações cerebrais relacionadas ao Alzheimer, além de sintomas novos, como dificuldades de memória, linguagem e pensamento. Esses problemas cognitivos podem ser perceptíveis para a pessoa afetada, sua família e amigos, mas não para outras pessoas, e podem não interferir em sua capacidade de realizar atividades diárias. Quando o cérebro é incapaz de reparar os danos e a morte de neurônios causados pela doença de Alzheimer, surgem problemas de memória, linguagem e habilidades de raciocínio
- **Demência Leve:** nesta fase, a maioria das pessoas consegue funcionar independentemente em muitas áreas, mas é provável precisar de ajuda em algumas atividades para maximizar a independência e permanecer segura
- **Demência Moderada:** nesta fase, que tende a ser a mais longa, as pessoas apresentam mais problemas de memória e linguagem, são mais propensas a ficarem confusas e têm mais dificuldade em realizar tarefas de várias etapas, como tomar banho e vestir-se. Eles podem ocasionalmente se tornar incontinentes e apresentar mudanças de personalidade e comportamento, como ansiedade e agitação. Eles também podem começar a ter dificuldades em reconhecer entes queridos
- **Demência Grave:** nesta fase, a capacidade das pessoas de se comunicarem verbalmente umas com as outras é muito diminuída e é provável precisarem de cuidados 24 horas. Os indivíduos ficam acamados devido a danos nas regiões cerebrais envolvidas no movimento. Ficar acamado os torna vulneráveis a complicações físicas, incluindo coágulos sanguíneos, infecções de pele e sepse, que desencadeiam inflamação em todo o corpo que pode resultar em falência de órgãos. Danos nas áreas do cérebro que controlam a deglutição dificultam a ingestão de alimentos e bebidas

2.1.5 Diagnóstico

Não há uma maneira simples de detectar a doença de Alzheimer. O diagnóstico requer um exame médico completo, a qual pode incluir: histórico médico familiar, exames de sangue, testes cognitivos e imagiologia cerebral podem ser usados para determinar a causa dos sintomas (ALZ, 2022).

Na prática, o diagnóstico é clínico, ou seja, depende da avaliação feita por um médico, que determinará as principais hipóteses para a origem do quadro com base na história do paciente e nos resultados de exames. Para descartar a possibilidade de outras doenças, exames de sangue e de imagem como tomografia ou, preferencialmente, ressonância magnética da cabeça devem ser realizados. Compõe parte da bateria de exames complementares uma avaliação aprofundada das funções cognitivas (ABRAZ, 2022).

Somente um exame microscópico do tecido cerebral do paciente após a morte pode fornecer um diagnóstico com certeza. Antes disso, esse exame não é recomendado por apresentar riscos para o paciente. A análise dos biomarcadores beta-amiloide e proteína tau, que estão sendo pesquisados para auxiliar no diagnóstico preciso, é uma novidade na pesquisa científica (ABRAZ, 2022).

As vantagens da ressonância magnética incluem sua ausência de radiação, alta resolução e baixa interferência, e seu uso em pacientes com doença de Alzheimer pode efetivamente identificar atrofia de várias estruturas cerebrais (WU *et al.*, 2022). Neste estudo serão utilizadas imagens de ressonância magnética para efetuar a classificação da doença de Alzheimer nas classes demente muito leve, demente leve, demente moderado e não demente.

Embora os médicos geralmente possam determinar se um paciente é portador de demência, muitas vezes pode ser difícil distinguir o tipo específico. Os erros de diagnóstico são mais frequentes na fase precoce. É crucial obter um diagnóstico adequado durante os estágios iniciais da doença, pois isso aumenta a probabilidade um paciente ser elegível para os tratamentos disponíveis, com o potencial de melhorar sua qualidade de vida (ALZ, 2022).

2.1.6 Tratamento

Atualmente não há cura para a doença de Alzheimer. Mesmo nos estágios mais graves da doença, os avanços médicos permitem que os pacientes vivam mais e com melhor qualidade de vida. A pesquisa avançou na compreensão dos mecanismos que causam a doença

e no desenvolvimento de medicamentos para seu tratamento. Os objetivos do tratamento são aliviar os sintomas, estabilizá-los ou, pelo menos, permitir uma progressão mais lenta da doença na maioria dos pacientes, permitindo que eles mantenham sua independência nas atividades diárias por períodos mais longos. Os tratamentos recomendados podem ser divididos entre os farmacológico e não farmacológico (ABRAZ, 2022).

Além disso, existem tratamentos não farmacológicos. Eles não alteram a biologia subjacente da doença e são frequentemente usados para manter ou melhorar a função cognitiva, a qualidade de vida e a capacidade de realizar atividades diárias. Os tratamentos não farmacológicos incluem terapia psicológica, terapias baseadas em música e estimulação cognitiva. Eles podem ser usados para diminuir sintomas comportamentais como depressão, apatia, estimulação, distúrbios do sono, agitação e agressividade (GAUGLER *et al.*, 2022).

2.2 Aprendizado de Máquina

O Aprendizado de Máquina é uma área de pesquisa da ciência da computação que envolve a indução de hipóteses ou funções, a partir de experiências passadas, para resolver um determinado problema de natureza preditiva ou descritiva (FACELI *et al.*, 2011). De acordo com Géron (2019) AM é a ciência (e a arte) da programação de computadores para que eles possam aprender com os dados.

Existem três tipos principais de aprendizado de máquina: supervisionado, não supervisionado e por reforço. No aprendizado supervisionado, o objetivo é prever os resultados com base em rótulos informados, classificação e regressão são tarefas típicas deste aprendizado. No aprendizado não supervisionado não são fornecidos rótulos, o objetivo é agrupar com base na similaridade e padrões. Já o aprendizado por reforço, o algoritmo não recebe a resposta correta, mas recebe um sinal de reforço, de recompensa ou punição. O algoritmo faz uma hipótese baseada nos exemplos e determina se essa hipótese foi boa ou ruim (LUDERMIR, 2021).

O AM tornou-se rapidamente o subcampo mais popular e bem-sucedido da inteligência artificial, uma tendência impulsionada pela disponibilidade de hardware mais rápido e conjuntos de dados maiores. Está intimamente relacionado com a estatística matemática, mas difere da estatística de várias maneiras importantes. Tende a lidar com conjuntos de dados grandes e complexos, onde a análise estatística clássica, como a análise bayesiana, é impraticável (CHOLLET, 2021).

2.2.1 *Aprendizado Supervisionado*

No aprendizado supervisionado, os dados de treinamento fornecidos ao modelo incluem os rótulos (as soluções desejadas). A classificação é uma tarefa típica do aprendizado supervisionado. O filtro de spam é um bom exemplo disso: ele é treinado com muitos exemplos de e-mails junto às classes (spam ou não spam) e deve aprender a classificar novos e-mails (GÉRON, 2019).

Prever um alvo de valor numérico é outra tarefa típica, como o preço de um carro a partir de um conjunto de características denominadas previsores. Esse tipo de tarefa é chamada de regressão. Para treinar o sistema, precisa-se fornecer muitos exemplos de carros, incluindo seus previsores e seus rótulos (ou seja, seus preços) (GÉRON, 2019). Este trabalho utilizará dos conceitos de aprendizado supervisionado e classificação, onde o objetivo é prever as múltiplas classes da doença de Alzheimer.

2.2.2 *Classificação*

Existem dois tipos de classificação: classificação binária, na qual os classificadores podem distinguir entre duas classes, e classificação multiclasse, na qual podem distinguir entre mais de duas classes. Alguns modelos, como os classificadores floresta aleatória e *Naive bayes*, conseguem lidar diretamente com múltiplas classes. Enquanto classificadores, como classificadores lineares e máquinas de vetores de suporte, são estritamente binários (GÉRON, 2019).

2.2.3 *Árvore de Decisão*

Árvore de decisão é um modelo versátil de aprendizado de máquina que pode ser executado em tarefas de classificação e regressão. É muito poderoso e capaz de moldar conjuntos complexos de dados (GÉRON, 2019). Consiste em estruturas baseadas em forma de árvore, que representam decisões e possíveis resultados com base em condições e atributos. As árvores de decisão particionam o conjunto de dados em diferentes caminhos, utilizando atributos como critérios de divisão, até chegar a uma decisão ou resultado (TECH, 2023).

A estrutura da árvore de decisão consiste em nós internos e folhas. Os nós internos representam testes em um atributo específico, enquanto as folhas representam as classes ou valores de saída resultantes da árvore. Neste modelo vários pontos de decisão são criados. Estes

pontos são os nós da árvore e em cada um deles o resultado da decisão será seguir por um caminho, ou por outro. Os caminhos existentes são os ramos.

2.2.4 *Árvores Extremamente Aleatórias*

Árvores extremamente aleatórias é um poderoso modelo de aprendizado de máquina baseado nas árvores de decisão. Uma variação das árvores de decisão, com uma aleatoriedade ainda mais forte. É basicamente um modelo floresta aleatória, porém mais aleatório ainda. Cada árvore construída no modelo irá escolher aleatoriamente qual limiar de atributo será usado em cada divisão (BRAINS, 2023).

2.2.5 *Floresta Aleatória*

A agregação de previsões de um conjunto de previsores, como classificadores ou regressores, muitas vezes obterá melhores previsões do que com o melhor previsor individual. Um conjunto de previsores é chamado de *ensemble*, assim, esta técnica é chamada *Ensemble Learning*, e um algoritmo de *Ensemble Learning* é chamado de *Ensemble method*. Floresta aleatória é um *ensemble*, ou seja, um conjunto de árvores de decisão (aleatórias), que tem como o propósito minimizar o sobreajuste de cada modelo individual de árvore gerado para os dados de entrada. Apesar da sua simplicidade, é um dos mais poderosos algoritmos de aprendizado de máquina disponível atualmente (GÉRON, 2019).

O modelo floresta aleatória é considerado um dos melhores, se não o melhor, algoritmo para lidar com dados não perceptivos atualmente. Em tarefas de classificação, a saída do algoritmo é a classe selecionada pela maioria das árvores. Já para tarefas de regressão é a previsão média das árvores individuais.

2.2.6 *KNN*

K-vizinhos mais próximos (KNN) é um dos modelos de classificação mais fundamentais e simples. Quando estimativas paramétricas confiáveis de densidade de probabilidade são desconhecidas ou difíceis de determinar, a análise discriminante deve ser realizada. Isso levou ao desenvolvimento da classificação K-vizinho mais próximo. Em 1951, Fix e Hodges introduziram um método não paramétrico para classificar padrões que desde então foi conhecido como KNN. O classificador KNN é comumente baseado na distância euclidiana entre uma

amostra de teste e as amostras de treinamento especificadas (PETERSON, 2009).

Portanto, o KNN tenta categorizar cada amostra de uma coleção de dados determinando a que distância ela está dos vizinhos mais próximos. Uma amostra será classificada nesta categoria se os vizinhos mais próximos forem predominantemente membros dessa classe. Quando o k é pequeno, a classificação se torna mais sensível às regiões próximas, talvez levando ao problema de sobreajuste. A classificação torna-se mais robusta e menos suscetível a ruído quando o k é grande, mas se o k for muito grande, pode haver um problema com subajuste.

2.2.7 SGD

Stochastic gradient descent / Gradiente descendente estocástico (SGD) é um modelo de aprendizado de máquina que trabalha muito bem com conjuntos de dados muito grandes. Isso se deve, em parte, ao fato de o SGD lidar independentemente com instâncias de treinamento, uma de cada vez, o que também torna o adequado para o aprendizado online. O gradiente descendente pega gradualmente os parâmetros, ou seja, ele ajusta iterativamente os parâmetros para minimizar a função custo no conjunto de treinamento (GÉRON, 2019).

O gradiente descendente é um modelo de otimização muito genérico capaz de encontrar ótimas soluções para uma ampla gama de problemas. o SGD escolhe uma instância aleatória no conjunto de treinamento em cada etapa e calcula os gradientes, baseado apenas nesta única instância. Isso torna o modelo muito mais rápido, pois tem poucos dados para manipular em cada iteração. Também permite treinar em grandes conjuntos de treinamento, uma vez que apenas uma instância precisa estar na memória a cada iteração (GÉRON, 2019).

2.2.8 Regressão Logística

A regressão logística é um classificador binário, que é normalmente empregada para estimar a probabilidade de uma instância pertencer a uma determinada classe. Se a probabilidade estimada for maior que 50%, então o modelo prevê que a instância pertence a essa classe, classe positiva, rotulada como 1, ou então prevê que não, classe negativa, rotulada 0. Um modelo de regressão logística calcula uma soma ponderada das características de entrada e gera a logística desse resultado. A logística é uma função sigmoide que mostra um número entre 0 e 1 (GÉRON, 2019).

2.2.9 *Naive Bayes*

Naive bayes é um modelo de aprendizado de máquina simples e eficiente baseado no teorema de *Bayes* com uma forte suposição de que os atributos são condicionalmente independentes dada a classe, ou seja, utiliza a probabilidade condicional para realizar a classificação de dados. É amplamente aplicado na prática, devido a sua eficiência computacional e muitos outros recursos desejáveis (WEBB *et al.*, 2010).

2.2.10 *SVM*

Uma Máquina de vetores de suporte (SVM) é um modelo de aprendizado de máquina muito poderoso e versátil, capaz de realizar classificações lineares ou não lineares, de regressão e até mesmo detecção de dados que se diferenciam drasticamente de todos os outros. É um dos modelos mais populares em aprendizado de máquina. As SVM são particularmente adequadas para a classificação de conjuntos de dados complexos (GÉRON, 2019).

O algoritmo de classificação SVM é o mais conhecido do grupo de algoritmos conhecidos como métodos do kernel. No início da década de 1990, no *Bell Labs*, Vladimir Vapnik e Corinna Cortes desenvolveram a atual fórmula. AS SVM visam resolver problemas de classificação por encontrar bons limites de decisão entre dois conjuntos de pontos pertencentes a duas categorias diferentes. Um limite de decisão pode ser pensado como uma linha ou superfície que separa seus dados de treinamento em dois espaços correspondentes a duas categorias. Para classificar novos pontos de dados, só precisa verificar em que lado do limite de decisão eles caem (CHOLLET, 2021).

Basicamente, uma SVM encontra uma linha de separação, usualmente conhecida como hiperplano, entre dados pertencentes a duas classes escolhidas arbitrariamente. Esta linha visa maximizar a distância entre os pontos que estão mais próximos de cada uma das classes. O principal objetivo da SVM é identificar o hiperplano ideal com as maiores margens para os pontos de dados. O modelo SVM é frequentemente usado em estudos para classificação de dados em diferentes grupos de doença de Alzheimer, ele será usado neste trabalho.

2.2.11 *GradientBoosting*

O modelo *GradientBoosting* é uma técnica de AM utilizada para problemas de classificação e regressão. É uma técnica *boosting*, incluída no grupo de classificadores *ensemble*.

Esses classificadores são métodos que usam uma combinação de resultados de preditores fracos, para produzir um melhor modelo preditivo. Quando os modelos de preditores fracos são utilizados individualmente possuem uma acurácia abaixo do esperado (SILVA, 2023).

Na técnica de *boosting*, cada classificador fraco é treinado com um conjunto de dados, sequencialmente e de uma maneira adaptativa, onde um modelo base depende dos anteriores, e no final são combinados de uma maneira determinística. O modelo *GradientBoosting* produz um modelo de previsão na forma de um conjunto de modelos de previsão fracos, geralmente árvores de decisão. É construído um modelo em etapas, como outros métodos de reforço, e os generaliza, permitindo a otimização de uma função de perda diferenciável arbitrária (SILVA, 2023).

2.2.12 XGBoost

XGBoost é um dos modelos de AM mais utilizados pelos cientistas de dados, com resultados superiores, principalmente em problemas de previsão envolvendo dados estruturados ou tabulares. Devido as suas características, consegue lidar eficientemente e com robustez com uma grande variedade de tipos diferentes de dados. *XGBoost* vem de *extreme gradient boosting*, e representa uma categoria de modelos baseada em árvores de decisão com *Gradient boosting* (MELO, 2023). O *XGBoost* é uma implementação de código aberto, conhecida e eficiente, do modelo baseado em árvores com aumento de gradiente (SAGEMAKER, 2023).

2.2.13 HistGradientBoosting

O modelo *HistGradientBoosting* é uma implementação alternativa do modelo *GradientBoosting* baseada em histograma. Esse estimador pode lidar com dados faltantes, ausentes, pois tem suporte integrado para valores ausentes (LUVSANDORJ, 2023).

2.2.14 Análise Discriminante Linear

A análise discriminante linear é um modelo de aprendizagem de máquina, que possui um limite de decisão linear. O modelo ajusta uma densidade gaussiana para cada classe, assumindo que todas as classes compartilham a mesma matriz de covariância (SKLEARN, 2023). A análise discriminante linear é uma ferramenta usada para classificação, redução de dimensão e visualização de dados. A meta em se aplicar técnicas de redução de dimensões é remover aquelas características redundantes e dependentes ao transformar características de um espaço

dimensional maior para um espaço dimensional menor (LAMFO, 2023).

2.2.15 Ridge

Ridge é um modelo de aprendizado de máquina baseado no método de regressão *Ridge* que converte os dados do rótulo em $[-1, 1]$ e resolve o problema com o método de regressão (DATATECHNOTES, 2023). No aprendizado de máquina, a classificação de *Ridge* é uma técnica usada para analisar modelos discriminantes lineares. É uma forma de regularização que penaliza os coeficientes do modelo para evitar o sobreajuste (KUMAR, 2023).

2.3 Balanceamento

Oversampling e *undersampling* são técnicas usadas na aprendizagem de máquina para lidar com desequilíbrios (desbalanceamento) na distribuição das classes em conjuntos de dados. Desequilíbrio de classes é quando há uma disparidade significativa entre o número de exemplos pertencentes a classe majoritária e a classe minoritária.

Oversampling consiste basicamente em aumentar a quantidade de exemplos da classe minoritária, seja por replicação de instâncias existentes ou criando sinteticamente novas instâncias. Já *undersampling* envolve a redução da quantidade de exemplos da classe majoritária, seja pela aplicação da seleção aleatória ou através da remoção de exemplos próximos a instâncias da classe minoritária (NETO, 2023).

2.3.1 SMOTE

A Técnica de sobreamostragem minoritária sintética (SMOTE) é uma técnica de *oversampling* utilizada para lidar com problemas de desequilíbrio das classes em conjuntos de dados. Para resolver este problema, a SMOTE aumenta o número de casos no conjunto de dados de um modo equilibrado. Criando novas instâncias das classes minoritárias existentes, sem mudar o número de classes majoritárias (LIKEBUPT, 2023).

As novas instâncias criadas pela SMOTE, não são apenas cópias de casos minoritários existentes, são utilizados exemplos do espaço de atributo para cada classe de destino e seus vizinhos mais próximos. Em seguida, são gerados novos exemplos que combinam atributos do caso de destino com recursos de seus vizinhos. Essa abordagem aumenta os atributos disponíveis para cada classe e torna os exemplos mais gerais (LIKEBUPT, 2023).

2.4 Aprendizado Profundo

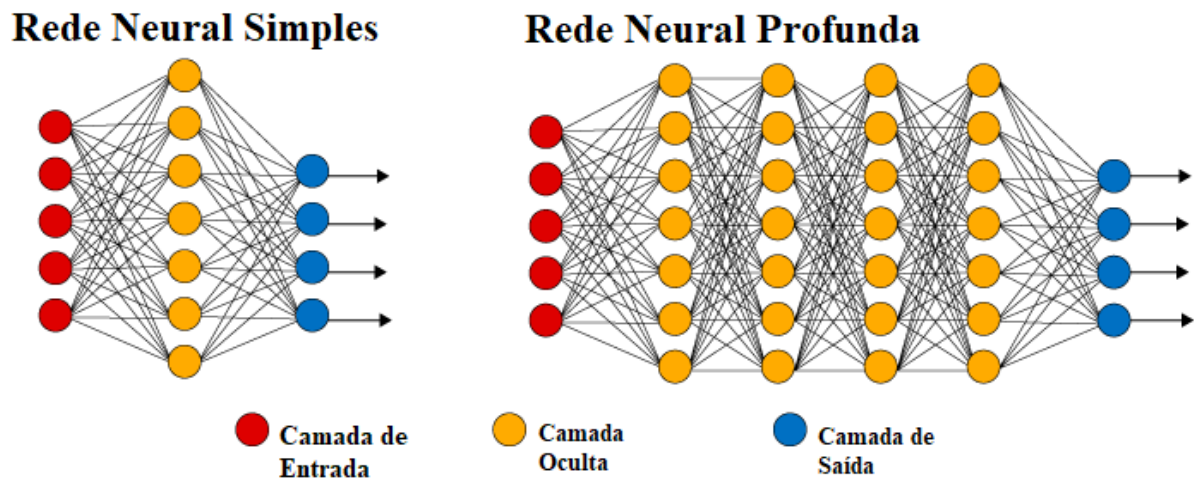
O aprendizado profundo é uma subárea do aprendizado de máquina, que emprega algoritmos para processar dados e imita o método de aprendizado que as pessoas usam para aprender tipos específicos de conhecimento, é um sistema baseado em neurônios. Pode ser pensado como uma forma automatizada de análise preditiva, que tem como ênfase sucessivas camadas com representações cada vez mais significativas. Esse sistema em que neurônios são organizados em camadas, com conexões entre neurônios de outras camadas consecutivas, foi chamado de rede neural artificial (CHOLLET, 2021). As redes neurais são adaptáveis, fortes e escaláveis, tornando-os perfeitas para lidar com tarefas de aprendizagem de máquina difíceis, como categorizar bilhões de imagens, conduzir serviços de reconhecimento de voz, recomendar os melhores vídeos para centenas de milhões de usuários diariamente (GÉRON, 2019).

O AP moderno geralmente envolve dezenas ou mesmo centenas de camadas sucessivas de representações e todas são aprendidas automaticamente da exposição a dados de treinamento. Enquanto isso, outras abordagens para o aprendizado de máquina tendem a concentrar-se no aprendizado de apenas uma ou duas camadas de representações dos dados (CHOLLET, 2021). Ele é mais adequado para questões desafiadoras, como reconhecimento de imagem e voz ou processamento de linguagem natural, contanto que tenha dados suficientes, capacidade de processamento e paciência (GÉRON, 2019). Ele será utilizado neste trabalho, visando melhores resultados.

2.4.1 *Redes Neurais*

Na aprendizagem profunda, essas representações em camadas são quase sempre aprendidas por modelos chamados redes neurais, estruturadas em camadas literais empilhadas umas sobre as outras. O termo rede neural é uma referência à neurobiologia, mas embora alguns dos conceitos centrais do AP tenham sido desenvolvidos em parte, inspirando-se na compreensão do cérebro, os modelos de aprendizado profundo não são modelos do cérebro. Não há evidências de que o cérebro implemente algo como os mecanismos de aprendizado usados nos modelos de aprendizado profundo (CHOLLET, 2021). A Figura 1 demonstra a representação de uma rede neural simples e uma rede neural profunda. As redes neurais são compostas por camada de entrada (primeira camada), camadas ocultas e camada de saída (última camada) (ACADEMY, 2022).

Figura 1 – Rede neural simples e rede neural profunda



Fonte: elaborado pelo autor (2023), imagem de Academy (2022).

2.4.2 Redes Neurais Convolucionais

As redes neurais convolucionais têm como inspiração o estudo do córtex visual do cérebro e são utilizadas no reconhecimento de imagens desde 1980. Nos últimos anos, graças ao aumento do poder computacional, a quantidade de dados de treinamento disponível, as redes neurais convolucionais conseguiram alcançar um desempenho sobre-humano em algumas tarefas visuais complexas. Elas fornecem serviços de pesquisa de imagens, carros autônomos, sistemas automáticos de classificação de vídeo e muito mais. Além disso, elas não estão restritas à percepção visual: sendo também bem-sucedidas em outras tarefas, como reconhecimento de voz ou processamento de linguagem natural (GÉRON, 2019).

Uma *Convolutional neural network* / Rede neural convolucional (CNN) é um algoritmo de aprendizado de máquina que pode captar uma imagem de entrada, atribuir importância (pesos) a vários aspectos / objetos da imagem e conseguir diferenciar um do outro. O pré-processamento exigido em uma rede neural convolucional é muito menor em comparação com outros algoritmos de classificação. Enquanto nos métodos primitivos os filtros são feitos à mão, com treinamento suficiente, elas conseguem aprender esses filtros / características. A arquitetura de uma CNN é análoga àquela do padrão de conectividade de neurônios no cérebro humano. Os neurônios individuais respondem a estímulos apenas em uma região restrita do campo visual conhecida como campo receptivo. Uma coleção desses campos se sobrepõe para cobrir toda a área visual (ACADEMY, 2022).

As redes neurais convolucionais são frequentemente usadas em estudos que utilizam aprendizagem profunda para efetuar classificação da doença de alzheimer, elas serão usadas

neste trabalho.

2.4.3 *Early Stopping*

Ao treinar redes neurais, o número de épocas de treinamento é muito importante, pois se forem usadas poucas épocas, o modelo poderá ter problemas de subajuste, ou seja, não aprenderá tudo o que poderia com os dados de treinamento. E se forem usadas muitas épocas, poderá ter o problema oposto, sobreajuste, ou seja, o modelo aprenderá demais. A *Early Stopping*, também conhecido como parada antecipada ou parada precoce, é uma técnica utilizada no treinamento de modelos para evitar o sobreajuste e determinar o momento ideal para parar o treinamento do modelo (ACADEMY, 2022). A *Early Stopping* interromperá o treinamento assim que as métricas de validação pararem de melhorar (CHOLLET, 2021).

2.5 Métricas de Avaliação

Esta seção descreve as métricas que serão utilizadas para avaliar os modelos propostos neste trabalho. Com a finalidade de estimar corretamente o desempenho dos modelos, serão utilizadas a acurácia, precisão, revocação e f1-score. O número de Verdadeiros positivos (VP), Verdadeiros negativos (VN), Falsos positivos (FP) e Falsos negativos (FN) influenciam nas métricas citadas. Essas métricas estão definidas abaixo.

- **VP:** é o número de verdadeiros positivos, ou seja, ocorre quando o modelo prevê corretamente a classe positiva. Por exemplo, quando o paciente está com a DA e o modelo previu corretamente que ele está com a DA.
- **VN:** é o número de verdadeiros negativos, ou seja, ocorre quando o modelo prevê corretamente a classe negativa. Exemplo: o paciente não estava com DA, e o modelo previu corretamente que ele não está.
- **FP:** é o número de falsos positivos, ou seja, ocorre quando o modelo prevê incorretamente a classe positiva. Por exemplo, o paciente não está com DA, mas o modelo disse que ele está.
- **FN:** é o número de falsos negativos, ou seja, ocorre quando o modelo prevê incorretamente a classe negativa. Exemplo: quando o paciente está com DA e o modelo previu incorretamente que ele não está.

2.5.1 Acurácia

A métrica acurácia demonstra a porcentagem de classificações corretas feitas sobre a quantidade total de classificações, definida pela equação 2.1. É considerada uma das métricas mais simples e importantes. Ela avalia simplesmente o percentual de acertos, ou seja, ela pode ser obtida pela razão entre a quantidade de acertos e o total de entradas.

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.1)$$

2.5.2 Precisão

A precisão é uma métrica determinada pela razão entre a quantidade de verdadeiros positivos e a soma de verdadeiros positivos com falsos positivos, definida pela equação 2.2, indica a porcentagem dos resultados do modelo preditos corretamente.

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (2.2)$$

2.5.3 Revocação

A revocação é uma métrica determinada pela razão entre a quantidade de verdadeiros positivos e a soma dos verdadeiros positivos com falsos negativos, definida pela equação 2.3, indica a porcentagem do total de previsões positivas corretas de todas as positivas que um modelo poderia ter feito.

$$\text{Revocação} = \frac{VP}{VP + FN} \quad (2.3)$$

2.5.4 F1-score

A métrica f1-score é determinada pela média harmônica entre a precisão e a revocação, definida pela equação 2.4. Enquanto a média regular trata igualmente todos os valores, a média harmônica dá muito mais peso aos valores mais baixos. Como resultado, o classificador só obterá uma pontuação f1-score alta se a revocação e a precisão forem altas. É uma medida da precisão de um teste onde o valor mais alto possível é 1. Isso indica precisão e recuperação perfeitas.

$$F1\text{-score} = \frac{2 * VP}{2 * VP + FP + FN} \quad (2.4)$$

3 TRABALHOS RELACIONADOS

Neste capítulo são apresentados os trabalhos que serviram como base de referência para o desenvolvimento desta pesquisa.

3.1 *Model Based on Deep Feature Extraction for Diagnosis of Alzheimer's Disease*

Em Silva *et al.* (2019), os pesquisadores propuseram um modelo para diagnóstico binário da doença de Alzheimer, baseado na extração de profundas características para a classificação por ressonância magnética. O objetivo principal do trabalho é classificar corretamente os pacientes que contêm doença de Alzheimer e que não contêm. O modelo consiste no uso de características geradas por uma CNN. Primeiramente, a rede neural foi pré-treinada para obter os melhores atributos. Depois os algoritmos de aprendizado de máquina supervisionados foram utilizados para a tarefa de classificação.

A base de dados utilizada pelos autores foi a *Minimal interval resonance imaging in Alzheimer's disease* (MIRIAD) para a realização do estudo. Ela consiste em imagens de ressonância magnética de 69 pessoas. E está dividida em duas classes: a primeira com 23 pacientes saudáveis e a segunda com 46 pacientes com a doença de Alzheimer. Foram selecionadas trinta fatias da região superior do cérebro para o aprendizado. A arquitetura convolucional foi projetada em três camadas convolucionais para extrair os melhores atributos da região selecionada. Em seguida, os atributos selecionados foram salvos em um vetor para o aprendizado e detecção de padrões.

Os dados foram particionados com o método de validação cruzada com k igual a 10, ou seja, foram efetuadas dez subdivisões (iguais). Depois foram treinados com os modelos floresta aleatória, SVM e KNN com diferentes parâmetros para avaliação. Os modelos foram avaliados com a métrica acurácia, tiveram como resultados 88,32%, 96,07% e 87,45%, para os algoritmos citados, respectivamente. Em uma análise comparativa realizada com outros trabalhos da literatura, ficou comprovado a eficiência e confiabilidade do modelo proposto.

O presente trabalho compartilha do mesmo objetivo de Silva *et al.* (2019), pois de maneira semelhante propõem prever corretamente, utilizando modelos de aprendizagem de máquina e aprendizagem profunda, eventos no domínio da doença de Alzheimer. A diferença deste trabalho para o dos autores será em relação à base de dados utilizada, o tipo de classificação: enquanto os autores efetuam uma classificação binária, este trabalho realizará uma classificação

multiclasse, e na quantidade de algoritmos avaliados. Além dos algoritmos de AM utilizados por Silva *et al.* (2019), o presente trabalho propõe testar outros algoritmos de AM e outras arquiteturas de redes neurais convolucionais, visando buscar melhores resultados.

3.2 Preclinical Stage Alzheimer's Disease Detection Using Magnetic Resonance Image Scans

Em Altay *et al.* (2021), os pesquisadores reforçam a preocupação com a doença de Alzheimer e a importância da detecção precoce, visando facilitar intervenções e tratamentos eficazes e em tempo hábil. Para resolver este problema desafiador, foram propostos dois mecanismos de atenção, um modelo de *Recurrent visual attention* / Atenção visual recorrente (RVA) e um modelo *Transformer*. Os dois modelos foram comparados com um modelo de linha de base, baseado em redes neurais convolucionais. Essa abordagem foi avaliada para diferenciar indivíduos em duas classes. A primeira classe com indivíduos rotulados com Alzheimer pré-clínico. E a segunda classe com indivíduos rotulados como outros. A classe outros inclui indivíduos saudáveis ou pacientes que sofrem de outros problemas de demência não relacionados com a doença de Alzheimer.

A base de dados Oasis-3 foi utilizado pelos autores para a realização do estudo. Ela consiste em imagens de ressonância magnética e tomografia por emissão de pósitrons. Com 1.098 indivíduos coletados em vários estudos ao longo de 15 anos. Existem 605 adultos cognitivamente normais e 493 indivíduos em diferentes estágios de declínio cognitivo. O conjunto de dados contém mais de 2.000 sessões de ressonância magnética. Esses dados incluem a idade, altura, peso e classificação de demência clínica.

Os modelos foram avaliados com as métricas f1-score, matriz de confusão, precisão, acurácia e taxa de falsos positivos. O modelo CNN obteve 83% de f1-score e 88,34% de acurácia. o modelo RVA obteve 87% de f1-score e 90,65% de acurácia. Já o modelo *Transformer* desenvolvido atingiu 90% de f1-score e 91,18% de acurácia na detecção pré-clínica.

Este trabalho diferencia-se do trabalho dos autores em relação: ao tipo de classificação, os autores efetuam uma classificação binária visando prever o estágio pré-clínico da doença de Alzheimer, esta pesquisa realizará uma classificação focada nas classes demente muito leve, demente leve, demente moderado e não demente.

3.3 DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images

Em Murugan *et al.* (2021), os pesquisadores propuseram um modelo que utiliza uma CNN para extrair as características discriminativas. A SMOTE foi utilizada para resolver o problema de desequilíbrio das classes da base de dados. Visando classificar os estágios da doença de Alzheimer com maior precisão, foi desenvolvido um modelo do zero. Os modelos foram avaliados e treinados para detectar os estágios do Alzheimer a partir de imagens de ressonâncias magnéticas.

A base de dados utilizada pelos autores para o desenvolvimento do modelo foi a *Alzheimer's Dataset (4 class of Images)*, que está disponível gratuitamente na plataforma *Kaggle*. Ela consiste em 6.400 imagens de ressonâncias magnéticas de quatro classes, são elas: demente muito leve, demente leve, demente moderado e não demente. O conjunto de dados tem um tamanho de imagem de 176 por 208 *pixels*.

O modelo desenvolvido pelos autores, *DEMNET*, obteve 95,23% na métrica acurácia e 97% na métrica área sob a curva, no conjunto de dados *Alzheimer's Dataset (4 class of Images)*. A base de dados da Iniciativa de neuroimagem da doença de Alzheimer (ADNI) foi utilizada para prever as classes da doença de Alzheimer, a fim de avaliar a eficácia do modelo proposto. O modelo obteve 84,83% de acurácia e de 95,62% área sob a curva.

O presente trabalho é similar ao do trabalho de Murugan *et al.* (2021), pois também classificará a doença de Alzheimer nas mesmas classes e utilizará a SMOTE para tratar o desbalanceamento da base de dados. A diferença deste trabalho para o dos pesquisadores será em relação à base de dados utilizada. Além disso, o presente trabalho propõe utilizar outros modelos de AM para a realização da classificação multiclasse.

3.4 Understanding Convolutional Neural Network's Behavior for Alzheimer's Disease on MRI

Em Islam *et al.* (2022), os pesquisadores propuseram uma nova arquitetura baseada em CNN para classificar a Doença de Alzheimer. O modelo consiste no uso de 5 camadas de convolução com 5 níveis de *pool* máximo. Depois disso, os autores tentaram simplificar a matriz multidimensional e, no final, usaram uma única camada densa para categorizar a doença de Alzheimer em seus quatro subtipos distintos.

A base de dados utilizada pelos autores para o desenvolvimento do modelo foi a *Alzheimer MRI Preprocessed*, que está disponível gratuitamente na plataforma *Kaggle*. Ela consiste em 6.400 imagens de ressonância magnética do cérebro humano. Ela está dividida em quatro classes, são elas: demente muito leve, demente leve, demente moderado e não demente. O conjunto de dados tem um tamanho de imagem de 128 por 128 *pixels*.

Os dados foram particionados em 90% das imagens usadas para treinamento e 10% usadas para teste. Além disso, 20% das imagens desses 90% dos conjuntos de dados de treinamento foram selecionadas para validação. Os conjuntos de dados de treinamento foram embaralhados para produzir imprevisibilidade e evitar problemas de superajuste. As métricas de avaliação utilizadas foram: acurácia, revocação, f1-score e matriz de confusão. Finalmente, foi usada uma camada densa com a função de ativação *softmax* para classificar as imagens. O modelo proposto foi avaliado e alcançou uma acurácia de 97%.

O presente trabalho utilizará a mesma base de dados utilizada por Islam *et al.* (2022). A diferença deste trabalho para o dos pesquisadores será nas arquiteturas das redes neurais convolucionais e o presente trabalho propõe utilizar a SMOTE, visando tratar o desbalanceamento dos dados.

3.5 Análise Comparativa

O Quadro 1 apresenta uma análise comparativa do presente trabalho e os trabalhos relacionados. Foram analisadas 4 características, são elas: base de dados, modelos, métricas e tipo de classificação.

Quadro 1 – Quadro comparativo entre os trabalhos relacionados e o trabalho proposto

Trabalhos	Silva et al. (2019)	Altay et al. (2021)	Murugan et al. (2021)	Islam et al. (2022)	Este trabalho
Base de Dados	MIRIAD	OASIS-3	Alzheimer's Dataset (4 class of Images) e ADNI	Alzheimer MRI Preprocessed	Alzheimer MRI Preprocessed
Modelos	KNN, SVM e Floresta Aleatória	CNN, RVA e Transformer	DEMNET	CNN	Árvore de Decisão, Árvores Extremamente Aleatórias, Floresta Aleatória, KNN, SGD, Regressão Logística, Naive Bayes, SVM, XGBoost, GradientBoosting, AdaBoost, HistGradientBoosting, Análise Discriminante Linear, Dummy, Ridge e CNN
Métricas	Acurácia, Sensibilidade, Especificidade e Área sob a curva	Acurácia, Precisão, F1-score, Matriz de Confusão e taxa de falsos positivos	Acurácia, Precisão, Revocação, F1-score, Área sob a curva e coeficiente de Kappa de Cohen	Acurácia, Revocação, F1-score e Matriz de Confusão	Acurácia, Precisão, Revocação e F1-score
Tipo de Classificação	Binária	Binária	Multiclasse	Multiclasse	Multiclasse

Fonte: elaborado pelo autor (2023).

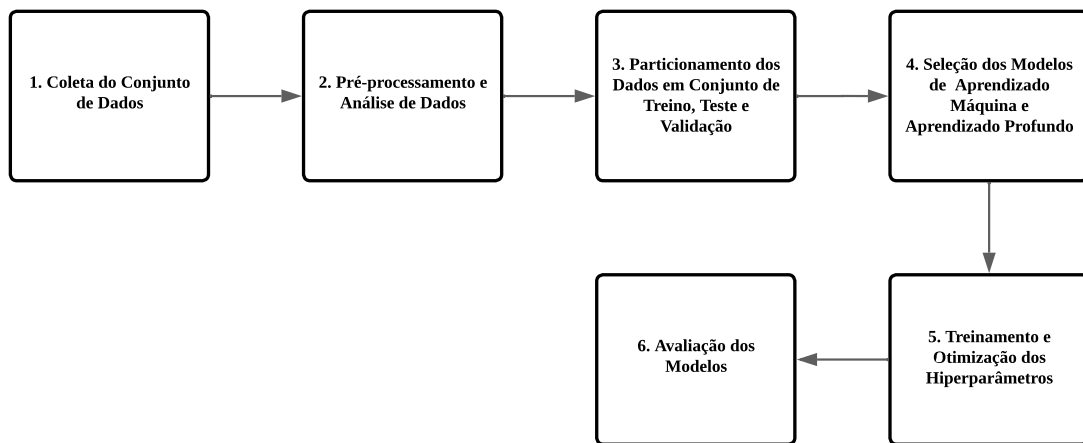
Todos os trabalhos utilizaram bases de dados no domínio da saúde e especificamente da doença de Alzheimer. As bases de dados contêm imagens de ressonância magnética. Todos

realizaram aprendizagem supervisionada para classificar Alzheimer em diferentes estágios da doença. Os dois primeiros realizaram classificação binária e os demais trabalhos, assim como o presente, realizaram classificação multiclasse. Todos os trabalhos utilizaram redes neurais convolucionais, uns para classificação, incluindo este trabalho, e outros para extração de atributos. O presente trabalho e o trabalho de Silva *et al.* (2019) utilizaram modelos de AM para a predição de Alzheimer. E por fim, todos os trabalhos utilizaram a métrica acurácia em conjunto com outras métricas.

4 PROCEDIMENTOS METODOLÓGICOS

A fim de alcançar os objetivos propostos neste trabalho, neste capítulo são apresentados um conjunto de etapas a serem seguidas. Conforme descrito nas próximas seções e no fluxograma, que pode ser visto na Figura 2.

Figura 2 – Fluxo de procedimentos metodológicos



Fonte: elaborado pelo autor (2022).

4.1 Coleta do Conjunto de Dados

O primeiro passo do procedimento metodológico consiste na triagem e obtenção de uma base de dados no domínio da saúde e especificamente da doença de Alzheimer, estruturada para um estudo de aprendizagem supervisionada, visando realizar a classificação multiclasse das amostras. Neste ponto, será realizada a busca de uma base dados de acesso gratuito, que contenha imagens de ressonância magnéticas do cérebro humano, rotuladas para a doença de Alzheimer.

4.2 Pré-processamento e análise de Dados

O segundo passo do procedimento metodológico consiste no pré-processamento dos dados e análise de dados. Será efetuada um conjunto de tarefas que envolvem preparação, organização e estruturação dos dados. Neste ponto, os dados serão tratados com o propósito de eliminar o ruído, que normalmente são os valores nulos, atípicos, fora de domínio, ausentes

e inconsistentes. Para escolher o melhor subconjunto de atributos, será feita uma seleção de atributos. Os atributos que serão mais úteis e significativos para resolver o problema são determinados nesta etapa.

Além disso, a engenharia de atributos poderá ser aplicada, os dados originais poderão ser alterados para formatos mais adequados. Nesta fase também será realizada a análise e visualização dos dados. Além disso, será efetuada a apresentação de informações em imagens ou tabelas. Para as imagens de ressonância magnética será aplicado o redimensionamento e normalização. Já para os rótulos serão convertidos de valores categóricos para valores numéricos.

4.3 Particionamento dos Dados em Conjunto de Treino, Teste e Validação

O terceiro passo do procedimento metodológico consiste no particionamento dos dados. Nesta etapa será efetuada a divisão dos dados em conjuntos de treino, testes e validação. O conjunto de treino será utilizado para treinar os modelos, o conjunto de validação será utilizado para validar os modelos e o conjunto de teste será utilizado para efetuar previsões, visando comprovar os modelos. Os dados serão particionados em 70% para o conjunto de treino, 20% para o conjunto de teste e 10% para o conjunto de validação.

4.4 Seleção dos Modelos de Aprendizado Máquina e Aprendizado Profundo

O quarto passo do procedimento metodológico consiste na escolha dos modelos que serão utilizados na implementação. Será efetuado um estudo exaustivo sobre os modelos de AM e AP para escolher os melhores modelos, visando alcançar os melhores resultados. A princípio os modelos de AM: KNN, SVM, floresta aleatória serão utilizados, estes modelos são os mais utilizados na classificação de Alzheimer em outros trabalhos acadêmicos e na plataforma *kaggle*. De AP será utilizado o modelo CNN, outro modelo bastante utilizado na classificação da doença de Alzheimer e para a classificação de imagens. Para o modelo CNN será proposto diversas arquiteturas, visando alcançar bons resultados.

4.5 Treinamento e Otimização dos Hiperparâmetros

O quinto passo do procedimento metodológico consiste no treinamento e otimização dos hiperparâmetros dos dados. Graças as etapas anteriores, os dados estão prontos para serem treinados. É nessa fase que os modelos selecionados na Seção 4.4 serão aplicados aos conjuntos

de treinamento, teste e validação gerados na Seção 4.3. Ainda nesta fase, será realizado a seleção dos melhores hiperparâmetros para cada modelo utilizando a técnica *GridSearch*, com a finalidade de encontrar uma ótima combinação de valores que permita um melhor desempenho.

4.6 Avaliação dos Modelos

O sexto passo do procedimento metodológico consiste na avaliação dos modelos. Nesta etapa serão aplicadas as métricas de classificação descritas na Seção 2.5 da fundamentação teórica. Será realizada uma análise quantitativa dos resultados dos modelos, a fim de realizar comparações. As métricas de avaliação incluirão acurácia, precisão, revocação e f1-score.

Por fim, será efetuada a interpretação e análise dos resultados obtidos. Este é o momento em que será feita as conclusões sobre os resultados obtidos. Uma análise das conclusões desta pesquisa será efetuada e comparada com o objetivo da pesquisa e os resultados esperados.

5 EXPERIMENTOS E RESULTADOS

Neste capítulo são apresentados os resultados obtidos a partir da execução dos procedimentos metodológicos. As ferramentas utilizadas para o desenvolvimento deste trabalho foram: a linguagem de programação Python, as bibliotecas *sklearn* para o pré-processamento das imagens e aplicação dos modelos de aprendizado de máquina, *imblearn* para a aplicação da SMOTE e *tensorflow* para a aplicação das redes neurais e pré-processamento. Os experimentos foram realizados em um notebook com processador Intel® Core i7-1165G7 e memória RAM com o total de 16 GB.

5.1 Coleta do Conjunto de Dados

Na etapa de coleta do conjunto de dados, prevista na Seção 4.1 dos procedimentos metodológicos, primeiro foi efetuada uma triagem das bases de dados estruturadas no domínio da saúde e especificamente da doença de Alzheimer para execução deste estudo. As pesquisas foram feitas em três locais. São eles:

- ADNI tem em vista desenvolver biomarcadores da doença de Alzheimer, avançar na compreensão da fisiopatologia, melhorar os métodos de diagnóstico para a detecção precoce. Outros objetivos incluem examinar a taxa de progressão para a doença de Alzheimer, bem como criar um banco de dados considerável de dados e imagens clínicas
- *kaggle* é uma plataforma e uma comunidade para aprendizado de ciência de dados. Por contar com muitos membros ativos, é considerada a maior comunidade nesta área. Além de destacar-se por apresentar competições premiadas,
- *Open access series of imaging studies (OASIS)* é um projeto que visa tornar conjuntos de dados de neuroimagem do cérebro disponível gratuitamente para a comunidade científica, com o intuito de facilitar futuras descobertas em neurociência básica e clínica

Foi solicitado o acesso às bases dados da ADNI e OASIS. ADNI concedeu o acesso a diversas bases de dados, OASIS não respondeu. A base definida para o estudo foi a *Alzheimer MRI Preprocessed*, que está disponível na plataforma *Kaggle*. Os fatores determinantes para essa escolha foram: trata-se de uma base livre, são dados coletados de vários repositórios públicos, alta quantidade de imagens (redimensionadas e de qualidade), foi utilizada em outros trabalhos e

por outros pesquisadores na plataforma *Kaggle*.

O conjunto de dados contém 6.400 imagens de ressonância magnética do cérebro e está dividida em quatro classes: 2240 imagens da classe demente muito leve, 896 da classe demente leve, 64 da classe demente moderado e 3200 da classe não demente. Todas as imagens são redimensionadas em 128 por 128 *pixels*.

5.2 Pré-processamento e Análise de Dados

Na etapa de pré-processamento e análise de dados, prevista na Seção 4.2 dos procedimentos metodológicos, foram realizados os seguintes procedimentos:

- **Conversão da escala de cores das imagens**
- **Aplicação da SMOTE**
- **Normalização**
- **Codificação de variáveis numéricas em variáveis categóricas**

Todos esses procedimentos tiveram como objetivo melhorar a qualidade dos dados e o desempenho dos modelos de predição.

5.2.1 Carregamento dos Dados

As imagens da base de dados obtida foram carregadas. Depois foi efetuado o redimensionamento das imagens para garantir que todas as imagens estivessem dimensionadas em 128 por 128 *pixels*. Em seguida, as imagens foram convertidas em vetores e salvas em uma lista de imagens, assim como seus rótulos. Por fim, as listas de imagens e rótulos foram convertidas em vetores e salvas nas variáveis X e y, respectivamente.

Para avaliar o desempenho dos modelos em mais de uma escala de coloração, as imagens foram carregadas novamente e convertidas da escala de cinza para a escala *RGB*. Todos os passos anteriores foram realizados novamente para as imagens na escala *RGB*.

5.2.2 Particionamento dos Dados em Conjunto de Treino, Teste e Validação

Na etapa de particionamento dos dados em conjunto de treino, teste e validação, prevista na Seção 4.3 dos procedimentos metodológicos, os dados foram particionados em 70% para o conjunto de treino, 20% para o conjunto de teste e 10% para o conjunto de validação. Visando tratar o desbalanceamento do conjunto de dados, foi aplicada a SMOTE no conjunto de

treino. O Quadro 2 demonstra a distribuição das classes nos conjuntos de dados.

Quadro 2 – Distribuição das classes nos conjuntos de dados

Classes	Treino	Treino (SMOTE)	Teste	Validação
Demente Muito Leve	1568	2240	448	224
Demente Leve	627	2240	179	90
Demente Moderado	45	2240	13	6
Não Demente	2240	2240	640	320
Total	4480	8960	1280	640

Fonte: elaborado pelo autor (2023).

5.2.3 Normalização e Codificação de Variáveis Categóricas

A técnica de normalização *Min-Max scaling* foi aplicada nas imagens, com objetivo de ajustar a escala de valores de *pixels* para a faixa entre 0 e 1. Também foi aplicada a conversão dos rótulos de valores categóricos para valores numéricos.

Para o treinamento dos modelos de aprendizado de máquina, as imagens foram transformadas em vetores de atributos com auxílio da função *flatten*, a qual é comumente utilizada para transformar vetores multidimensionais em vetores unidimensionais. Já para os modelos de aprendizado profundo, os rótulos numéricos foram convertidos para vetores codificados.

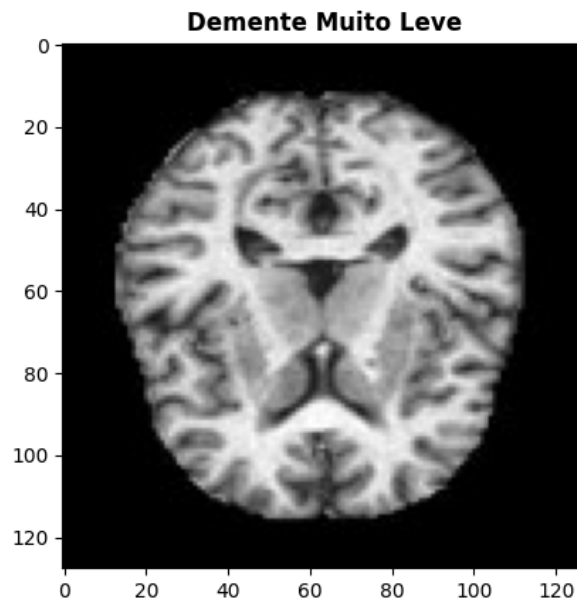
5.2.4 Análise de Dados

A base dados consiste em imagens de ressonâncias magnéticas e seus respectivos rótulos. Todas as imagens estão no formato JPG e redimensionadas em 128 por 128 *pixels*. A limpeza para eliminar ruídos não foi necessária, pois não existem valores nulos, atípicos, fora de domínio, ausentes e inconsistentes. As figuras 3, 4, 5 e 6 são imagens de ressonância magnética que foram plotadas para a visualização.

5.3 Seleção dos Modelos de Aprendizado de Máquina e Aprendizado Profundo

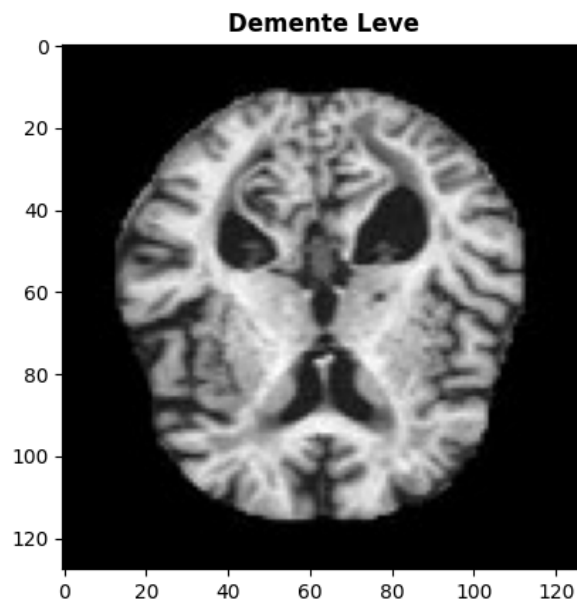
Na etapa de seleção dos modelos de aprendizado de máquina e aprendizado profundo, prevista na Seção 4.4 dos procedimentos metodológicos, foram selecionados 15 modelos de aprendizagem de máquina e 4 arquiteturas de aprendizagem profunda. De AM foram selecionados os modelos: árvore de decisão, árvores extremamente aleatórias, floresta aleatória, KNN, SGD, regressão logística, *Naive bayes* (multinomial e gaussiano), SVM (RBF, linear, poly,

Figura 3 – Imagem de ressonância magnética da classe demente muito leve



Fonte: elaborado pelo autor (2023), imagens da base dados *Alzheimer MRI Preprocessed Dataset*.

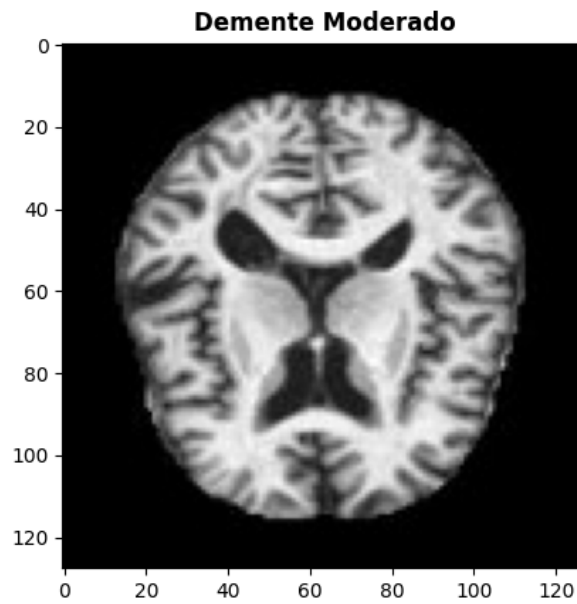
Figura 4 – Imagem de ressonância magnética da classe demente leve



Fonte: elaborado pelo autor (2023), imagens da base dados *Alzheimer MRI Preprocessed Dataset*.

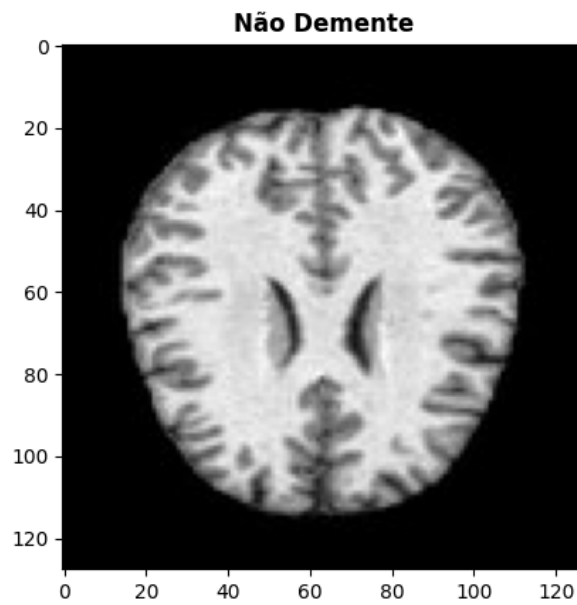
sigmoid), *XGBoost*, *GradientBoosting*, *adaBoost*, *HistGradientBoosting*, análise discriminante linear, *dummy*, *ridge*. De AP foi selecionado o modelo rede neural convolucional.

Figura 5 – Imagem de ressonância magnética da classe demente moderado



Fonte: elaborado pelo autor (2023), imagens da base dados *Alzheimer MRI Preprocessed Dataset*.

Figura 6 – Imagem de ressonância magnética da classe não demente



Fonte: elaborado pelo autor (2023), imagens da base dados *Alzheimer MRI Preprocessed Dataset*.

5.4 Treinamento e Otimização dos Hiperparâmetros

Na etapa de treinamento e otimização dos hiperparâmetros, prevista na Seção 4.5 dos procedimentos metodológicos, os modelos de aprendizado de máquina selecionados foram treinados com o conjunto de treinamento e realizadas as previsões com o conjunto de teste. Já os modelos de aprendizado profundo foram treinados com o conjunto de treinamento e com o conjunto de validação.

Todos os modelos de AP foram compilados com o otimizador *adam*, com perda como *categorical_crossentropy* e utilizando as métricas acurácia, precisão, revocação e f1-score. No treinamento foi configurado 30 épocas e a técnica *Early Stopping* foi usada, para evitar o sobreajuste e melhorar o desempenho geral dos modelos.

5.4.1 Arquitetura das Redes Neurais

As Figuras 7, 8, 9, 10 apresentam as quatro arquiteturas das redes neurais proposta para este trabalho. Pode-se observar que a entrada da primeira camada das quatro arquiteturas apresenta (None, 128, 128, 1) a qual é a especificação das dimensões das imagens, 128 por 128 pixels, o 1 é a coloração, que está na escala de cinza. Para as imagens na coloração RGB, a primeira camada apresenta uma sutil diferença, ela apresenta (None, 128, 128, 3), o 3 é da coloração RGB.

A arquitetura 01, Figura 7, apresenta três camadas de convolução com 25, 64 e 64 filtros, respectivamente, com função de ativação *reLU*. Duas camadas de *pooling* com janelas de tamanho (2, 2). Uma camada (*flatten*) e duas camadas densas, a primeira com 64 neurônios e função de ativação *reLU* e a última, a camada de saída, com 4 neurônios e função de ativação *softmax*.

A arquitetura 02, Figura 8, apresenta três camadas de convolução com 32, 64 e 64 filtros, respectivamente, com função de ativação *reLU*. Duas camadas de *pooling* com janelas de tamanho (2, 2). Uma camada (*flatten*) e duas camadas densas, a primeira com 64 neurônios e função de ativação *reLU* e a última, a camada de saída, com 4 neurônios e função de ativação *softmax*.

A arquitetura 03, Figura 9, apresenta três camadas de convolução com 64, 32 e 32 filtros, respectivamente, com função de ativação *reLU*. Duas camadas de *pooling* com janelas de tamanho (2, 2). Uma camada (*flatten*) e três camadas densas, as duas primeiras com 100 e

50 neurônios, respectivamente e função de ativação *reLU* e a última, a camada de saída, com 4 neurônios e função de ativação *softmax*.

A arquitetura 04, Figura 10, apresenta três camadas de convolução com 32, 64 e 128 filtros, respectivamente, com função de ativação *reLU*. três camadas de *pooling* com janelas de tamanho (2, 2). Uma camada (*flatten*) e duas camadas densas, a primeira com 512 neurônios e função de ativação *reLU* e a última, a camada de saída, com 4 neurônios e função de ativação *softmax*. Entre as camadas densas existe uma camada de *dropout* com taxa de 0.5, ou seja, durante o treino, metade dos neurônios dessa camada serão aleatoriamente desativados, visando evitar o sobreajuste.

Para todas as quatro arquiteturas, a última camada é a camada de saída, a qual é uma camada densa, com 4 neurônios, correspondentes as classes a serem preditas. Nesta camada é utilizada uma função de ativação *softmax*, utilizada para normalizar as saídas em uma distribuição de probabilidade sobre as classes.

5.5 Avaliação dos Modelos

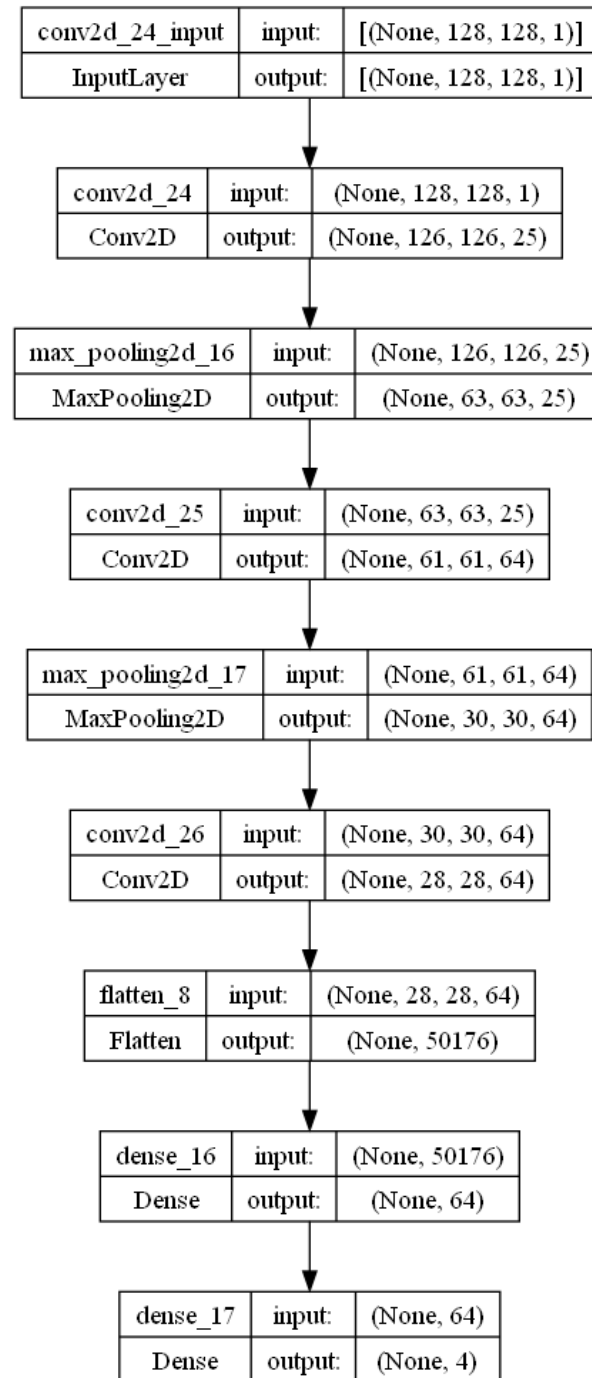
Na etapa de avaliação dos modelos, prevista na Seção 4.6 dos procedimentos metodológicos, os modelos foram avaliados com base nas métricas: acurácia, precisão, revocação e f1-score. Para efeito comparativo, os modelos foram executados primeiro para as imagens na escala de cinza e uma segunda vez para as imagens na escala *RGB*.

5.5.1 Cenários de Experimentação

Definidas as estratégias de treinamento, os cenários de experimentação foram criados para o treinamento e avaliação dos modelos. Os objetivos específicos de cada experimento são detalhados abaixo.

- Experimento 1: Avaliar o desempenho dos modelos de aprendizado de máquina com hiperparâmetros padrão
- Experimento 2: Avaliar o desempenho dos modelos utilizando a validação cruzada
- Experimento 3: Avaliar o impacto do balanceamento
- Experimento 4: Avaliar o desempenho dos modelos utilizando ajuste de hiperparâmetros

Figura 7 – Arquitetura rede neural convulcional 01



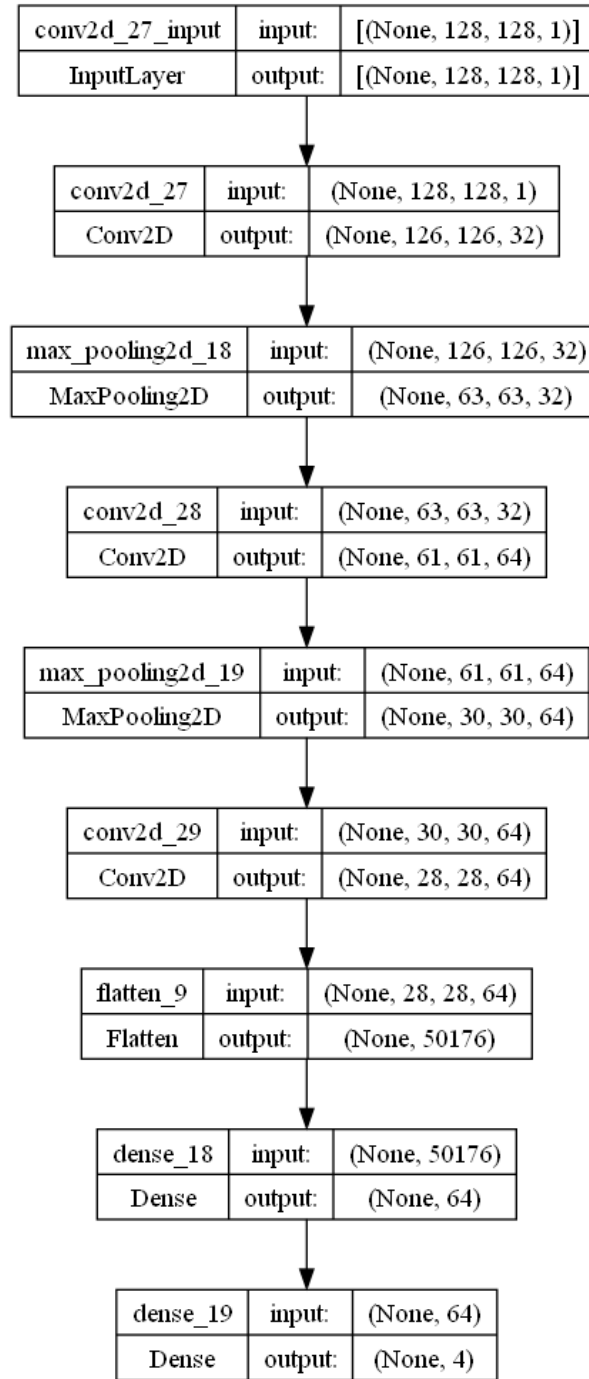
Fonte: elaborado pelo autor (2023).

- Experimento 5: Avaliar o desempenho dos modelos de aprendizado profundo

5.5.1.1 Experimento 1

Este experimento teve como objetivo avaliar o desempenho dos modelos de aprendizado de máquina utilizando os hiperparâmetros padrão. A tabela 1 detalha os resultados deste

Figura 8 – Arquitetura rede neural convulcional 02



Fonte: elaborado pelo autor (2023).

experimento. Os resultados destacados em **negrito** indicam os maiores valores para cada métrica.

Para o experimento 1, o modelo SVM com *kernel* linear alcançou **97,03%**, em todas as métricas, os melhores resultados em comparação aos outros modelos. Em segundo lugar ficou o modelo *Ridge*, que se saiu melhor com as imagens na escala de cinza do que na escala *RGB*. Depois em terceiro lugar ficou o *HistGradientBoosting* que obteve resultados iguais para as

Tabela 1 – Experimento 1, avaliação dos modelos com hiperparâmetros padrão

Modelo	Modo	Acurácia	Precisão	Revocação	F1-score
Árvore de Decisão	Escala de cinza	70,00%	69,69%	70,00%	69,79%
	RGB	68,44%	68,21%	68,44%	68,30%
Árvores Extremamente Aleatórias	Escala de cinza	64,45%	64,73%	64,45%	64,57%
	RGB	67,42%	67,37%	67,42%	67,38%
Floresta Aleatória	Escala de cinza	88,98%	89,53%	88,98%	88,63%
	RGB	89,38%	89,93%	89,38%	88,85%
K-vizinhos mais próximos	Escala de cinza	95,55%	95,57%	95,55%	95,53%
	RGB	95,55%	95,57%	95,55%	95,53%
SGD	Escala de cinza	92,19%	92,30%	92,19%	92,14%
	RGB	92,89%	93,03%	92,89%	92,86%
Regressão Logística	Escala de cinza	88,12%	88,17%	88,12%	88,07%
	RGB	88,28%	88,33%	88,28%	88,24%
Naive Bayes Multinomial	Escala de cinza	48,83%	49,98%	48,83%	48,55%
	RGB	48,67%	50,09%	48,67%	48,50%
Naive Bayes Gaussiano	Escala de cinza	52,34%	55,27%	52,34%	52,69%
	RGB	52,34%	55,31%	52,34%	52,70%
SVC (RBF)	Escala de cinza	74,53%	75,03%	74,53%	72,36%
	RGB	74,53%	75,03%	74,53%	72,36%
SVC (Linear)	Escala de cinza	97,03%	97,03%	97,03%	97,03%
	RGB	97,03%	97,03%	97,03%	97,03%
SVC (poly)	Escala de cinza	91,02%	91,19%	91,02%	90,91%
	RGB	91,02%	91,19%	91,02%	90,91%
SVC (sigmoid)	Escala de cinza	35,63%	35,35%	35,63%	35,48%
	RGB	35,63%	35,35%	35,63%	35,48%
XGB	Escala de cinza	95,78%	95,86%	95,78%	95,73%
	RGB	95,78%	95,86%	95,78%	95,73%
GradientBoosting	Escala de cinza	87,58%	87,91%	87,58%	87,28%
	RGB	87,58%	87,94%	87,58%	87,32%
AdaBoost	Escala de cinza	57,89%	56,12%	57,89%	55,94%
	RGB	57,89%	56,12%	57,89%	55,94%
HistGradientBoosting	Escala de cinza	96,64%	96,71%	96,64%	96,59%
	RGB	96,64%	96,71%	96,64%	96,59%
Análise Discriminante Linear	Escala de cinza	83,52%	83,68%	83,52%	83,56%
	RGB	83,52%	83,68%	83,52%	83,56%
Dummy	Escala de cinza	50,00%	25,00%	50,00%	33,33%
	RGB	50,00%	25,00%	50,00%	33,33%
Ridge	Escala de cinza	96,95%	96,95%	96,95%	96,94%
	RGB	96,88%	96,87%	96,88%	96,86%

Fonte: elaborado pelo autor (2023).

imagens nas duas escalas.

5.5.1.2 Experimento 2

Este experimento teve como objetivo avaliar o desempenho dos modelos de aprendizado de máquina utilizando a validação cruzada com k igual a 5, ou seja, significa que foram efetuadas cinco subdivisões iguais. A tabela 2 detalha os resultados deste experimento. Os

resultados destacados em negrito indicam os maiores valores para cada métrica.

Tabela 2 – Experimento 2, avaliação dos modelos utilizando a validação cruzada

Modelo	Acurácia	Precisão	Revocação	F1-score
Árvore de Decisão	53,98%	53,83%	53,98%	53,88%
Árvores Extremamente Aleatórias	51,88%	51,99%	51,88%	51,87%
Floresta Aleatória	68,52%	69,73%	68,52%	65,70%
K-vizinhos mais próximos	64,53%	64,36%	64,53%	62,70%
SGD	69,53%	70,10%	69,53%	69,69%
Regressão Logística	75,23%	75,18%	75,23%	74,95%
Naive Bayes Multinomial	52,50%	48,79%	52,50%	48,79%
Naive Bayes Gaussiano	47,89%	49,95%	47,89%	46,45%
SVC (RBF)	59,38%	49,66%	59,38%	53,86%
SVC (Linear)	76,80%	76,73%	76,80%	76,69%
SVC (poly)	70,47%	71,23%	70,47%	69,49%
SVC (sigmoid)	37,34%	27,94%	37,34%	30,91%
XGB	95,78%	95,86%	95,78%	95,73%
GradientBoosting	87,58%	87,91%	87,58%	87,28%
AdaBoost	57,89%	56,12%	57,89%	55,94%
HistGradientBoosting	96,64%	96,71%	96,74%	96,59%
Análise Discriminante Linear	83,52%	83,68%	83,52%	83,56%
Dummy	50,00%	25,00%	50,00%	33,33%
Ridge	96,95%	96,95%	96,95%	96,94%

Para este experimento, O modelo *Ridge* alcançou 96,95% em todas as métricas, os melhores resultados em comparação aos outros modelos. Depois tiveram bons resultados, os modelos *HistGradientBoosting*, *XGBoost* e *GradientBoosting*.

O trabalho de Silva *et al.* (2019) obteve como resultados 88,32% para o modelo floresta aleatória, 96,07% para o modelo SVM e 87,45% para o modelo KNN na métrica acurácia. Para os mesmo modelos este trabalho obteve como resultados 64,52%, 64,53% e 70,47%, para os modelos citados, respectivamente. Vale ressaltar que Silva *et al.* (2019) utilizou k igual a 10 e que os modelos deste experimento estava com hiperparâmetros padrão. O modelo floresta aleatória no experimento 4, resultados na Tabela 4, onde o k era igual a 2, obteve 91,56% de acurácia, resultado superior aos 88,32% de Silva *et al.* (2019).

5.5.1.3 Experimento 3

Este experimento teve como objetivo avaliar o impacto do balanceamento no desempenho dos modelos de aprendizado de máquina, a técnica de balanceamento SMOTE foi

utilizada. A tabela 3 detalha os resultados deste experimento. Os resultados destacados em negrito indicam os maiores valores para cada métrica.

Tabela 3 – Experimento 3, avaliação dos modelos utilizando a SMOTE e com hiperparâmetros padrão

Modelo	Modo	Acurácia	Precisão	Revocação	F1-score
Árvore de Decisão	Escala de cinza	68,13%	68,82%	68,13%	68,38%
	RGB	68,13%	68,42%	68,13%	68,23%
Árvores Extremamente Aleatórias	Escala de cinza	64,77%	65,91%	64,77%	65,16%
	RGB	63,20%	64,57%	63,20%	63,71%
Floresta Aleatória	Escala de cinza	93,13%	93,21%	93,13%	93,11%
	RGB	93,67%	93,69%	93,67%	93,66%
K-vizinhos mais próximos	Escala de cinza	92,73%	94,12%	92,73%	92,97%
	RGB	92,73%	94,12%	92,73%	92,97%
SGD	Escala de cinza	90,55%	91,20%	90,55%	90,44%
	RGB	91,87%	91,88%	91,87%	91,86%
Regressão Logística	Escala de cinza	88,12%	88,14%	88,12%	88,13%
	RGB	86,02%	86,06%	86,02%	86,02%
Naive Bayes Multinomial	Escala de cinza	47,58%	51,42%	47,58%	47,91%
	RGB	47,58%	51,42%	47,58%	47,91%
Naive Bayes Gaussiano	Escala de cinza	53,98%	55,41%	53,98%	53,97%
	RGB	53,98%	55,41%	53,98%	53,97%
SVC (RBF)	Escala de cinza	82,03%	82,71%	82,03%	82,13%
	RGB	82,03%	82,71%	82,03%	82,13%
SVC (Linear)	Escala de cinza	97,03%	97,03%	97,03%	97,03%
	RGB	97,03%	97,03%	97,03%	97,03%
SVC (poly)	Escala de cinza	93,20%	93,21%	93,20%	93,20%
	RGB	93,20%	93,21%	93,20%	93,20%
SVC (sigmoid)	Escala de cinza	22,73%	31,94%	22,73%	20,74%
	RGB	22,73%	31,94%	22,73%	20,74%
XGB	Escala de cinza	97,03%	97,08%	97,03%	97,02%
	RGB	97,03%	97,08%	97,03%	97,02%
GradientBoosting	Escala de cinza	84,61%	84,61%	84,61%	84,61%
	RGB	84,61%	84,61%	84,61%	84,61%
AdaBoost	Escala de cinza	50,55%	53,17%	50,55%	50,82%
	RGB	50,55%	53,17%	50,55%	50,82%
HistGradientBoosting	Escala de cinza	97,66%	97,69%	97,66%	97,64%
	RGB	97,66%	97,69%	97,66%	97,64%
Análise Discriminante Linear	Escala de cinza	81,80%	82,08%	81,80%	81,88%
	RGB	81,80%	82,08%	81,80%	81,88%
Dummy	Escala de cinza	13,98%	1,96%	13,98%	3,43%
	RGB	13,98%	1,96%	13,98%	3,43%
Ridge	Escala de cinza	97,03%	97,03%	97,03%	97,02%
	RGB	96,88%	96,87%	96,88%	96,86%

Fonte: elaborado pelo autor (2023).

Para este experimento, o modelo *HistGradientBoosting* alcançou 97,66% nas métricas acurácia, precisão e revocação e 97,64% na métrica f1-score, os melhores resultados em comparação aos outros modelos. Depois tiveram bons resultados: *Ridge*, *XGBoost*, SVM com

kernel linear e *poly*. Os modelos floresta aleatória e o *HistGradientBoosting* obtiveram melhores resultados neste experimento em comparação com os resultados do experimento 1.

5.5.1.4 Experimento 4

Este experimento teve como objetivo avaliar o desempenho dos modelos de aprendizado de máquina utilizado a otimização de hiperparâmetros com a técnica *GridSearch*. A tabela 4 detalha os resultados deste experimento. As tabelas 5, 6, 7 e 8 detalham os hiperparâmetros utilizados pelos modelos. Os resultados destacados em negrito indicam os maiores valores para cada métrica.

Tabela 4 – Experimento 4, avaliação dos modelos utilizando a técnica *gridSearch*

Modelo	Hiperparâmetro	Acurácia	Precisão	Revocação	F1-score
Árvore de decisão	Não	70,00%	69,69%	70,00%	69,79%
	Sim	70,55%	70,48%	70,55%	70,46%
Árvores extras	Não	64,45%	64,73%	64,45%	64,57%
	Sim	65,39%	65,25%	65,39%	65,31%
Floresta Aleatória	Não	88,98%	89,53%	88,98%	88,63%
	Sim	91,56%	92,05%	91,56%	91,38%
AdaBoost	Não	57,89%	56,12%	57,89%	55,94%
	Sim	55,94%	55,44%	55,94%	55,50%

Fonte: elaborado pelo autor (2023).

Para este experimento, todos os modelos testados com os ajustes de hiperparâmetros alcançaram melhores resultados em comparação aos modelos com hiperparâmetros padrão. Não foi possível efetuar os testes para todos os modelos, devido ao poder computacional limitado.

Tabela 5 – Hiperparâmetros do modelo árvore de decisão

Árvore de Decisão		
Hiperparâmetros	Avaliados	Melhores
max_depth	None, 5, 10, 15	15
min_samples_split	2, 5, 10	2
min_samples_leaf	1, 2, 3, 4	1
criterion	gini, entropy, log_loss	entropy

Fonte: elaborado pelo autor (2023).

5.5.1.5 Experimento 5

Este experimento visou avaliar o desempenho dos modelos de aprendizado profundo. A Tabela 9 detalha a perda de cada modelo e os resultados das métricas. Os resultados destacados

Tabela 6 – Hiperparâmetros do modelo árvores extras

<i>Árvores Extras</i>		
Hiperparâmetros	Avaliados	Melhores
min_samples_leaf	1, 2, 3, 4	1
criterion	gini, entropy, log_loss	entropy
max_features	auto, sqrt, log2	auto

Fonte: elaborado pelo autor (2023).

Tabela 7 – Hiperparâmetros do modelo floresta aleatória

<i>Floresta Aleatória</i>		
Hiperparâmetros	Avaliados	Melhores
n_estimators	100, 150, 200, 250, 300	250
min_samples_leaf	1, 2, 3, 4, 5	1
criterion	gini, entropy, log_loss	gini
max_features	auto, sqrt, log2	auto

Fonte: elaborado pelo autor (2023).

Tabela 8 – Hiperparâmetros do modelo *adaBoost*

<i>AdaBoost</i>		
Hiperparâmetros	Avaliados	Melhores
n_estimators	100, 200, 300, 400, 500	400
learning_rate	1, 2, 3, 4, 5	1

Fonte: elaborado pelo autor (2023).

em negrito indicam os maiores valores para cada métrica e o menor valor de perda.

Tabela 9 – Experimento 5, avaliação dos modelos de aprendizado profundo

Modelo	Modo	Acurácia	Precisão	Revocação	F1-score	Perda
Modelo 01	Escala de cinza	97,66%	97,66	97,66%	97,96%	7,81%
	Escala de cinza (SMOTE)	97,81%	97,81	97,81%	97,98%	7,28%
	RGB	98,05%	98,28	98,05%	98,29%	8,55%
	RGB (SMOTE)	96,95%	97,03%	96,95%	97,64%	9,81%
Modelo 02	Escala de cinza	97,89%	97,89%	97,89%	98,13%	9,65%
	Escala de cinza (SMOTE)	96,88%	96,88%	96,88%	97,18%	13,43%
	RGB	98,05%	98,05%	97,97%	98,05%	9,11%
	RGB (SMOTE)	97,97%	97,97%	97,97%	98,19%	10,54%
Modelo 03	Escala de cinza	96,95%	97,03%	96,95%	97,36%	10,73%
	Escala de cinza (SMOTE)	96,88%	97,02	96,72%	97,20%	10,54%
	RGB	97,11%	97,11%	97,11%	97,22%	13,19%
	RGB (SMOTE)	96,25%	96,24%	96,09%	96,73%	15,15%
Modelo 04	Escala de cinza	97,73%	97,73%	97,58%	98,02%	9,74%
	Escala de cinza (SMOTE)	97,50%	97,50%	97,50%	97,53%	8,03%
	RGB	96,88%	96,95%	96,88%	97,08%	9,97%
	RGB (SMOTE)	96,56%	96,63%	96,41%	97,07%	10,57%

Fonte: elaborado pelo autor (2023).

O modelo 01 utilizando as imagens na escala *RGB* obteve 98,05% na métrica acurácia, 98,28% na métrica precisão, 98,05% na métrica revocação e 98,29% na métrica f1-score, foi o modelo que alcançou os melhores resultados em comparação com os outros três modelos deste experimento. O modelo 01 utilizando as imagens na escala de cinza e a SMOTE obteve a menor

perda. O modelo 02 utilizando as imagens na escala *RGB* também obteve 98,05% na métrica acurácia, porém obteve resultados inferiores nas outras métricas em comparação ao modelo 01.

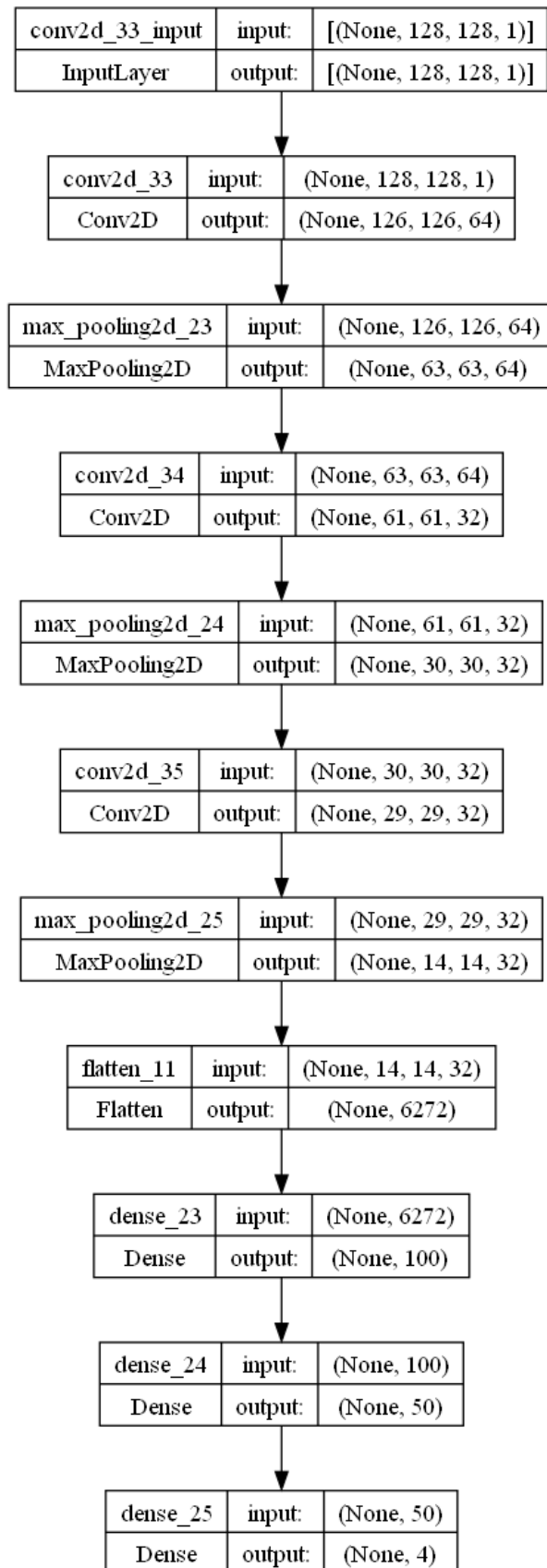
O trabalho de Islam *et al.* (2022), que utilizou a mesma base dados que este trabalho e também avaliou uma CNN, obteve 97% na métrica acurácia. Os modelos 01, 02 e 04 obtiveram resultados superiores a 97% na métrica acurácia, resultados melhores em comparação com o trabalho de Islam *et al.* (2022).

5.6 Interpretação e Análise dos Resultados

Os modelos de aprendizado profundo obtiveram os melhores resultados em comparação aos modelos de aprendizado de máquina. Já o modelo 01 foi o modelo que alcançou os melhores resultados em comparação com os outros três modelos de AP testados e em comparação a todos os modelos de AM testados. Portanto, com base nos experimentos, o modelo 01 é o melhor modelo para a realização do objetivo principal deste trabalho, ou seja, é o modelo que melhor realiza a predição de diagnósticos da doença de Alzheimer nas classes demente muito leve, demente leve, demente moderado e não demente, usando imagens de ressonância magnética.

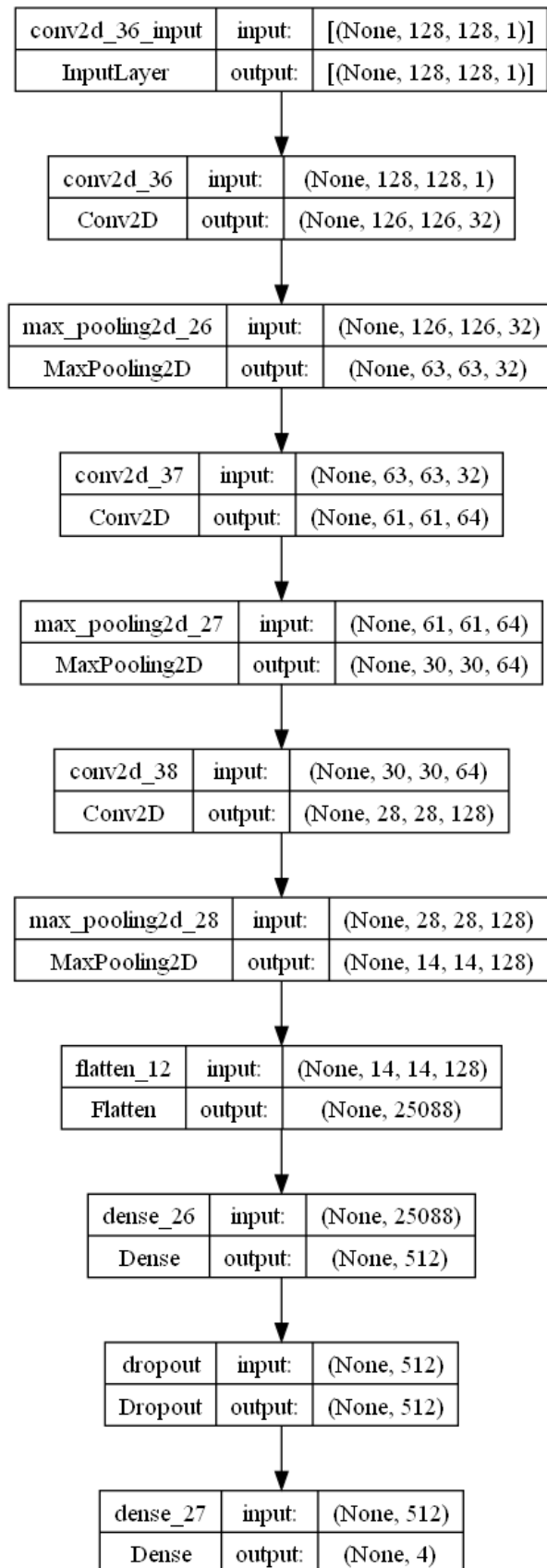
Por fim, conclui-se que os resultados obtidos nesta pesquisa atenderam as expectativas dos resultados esperados. E que todos os objetivos propostos, foram realizados. Alguns experimentos foram dificultados devido ao *hardware* utilizado nesta pesquisa, todavia medidas para contornar este problema foram realizadas, como, por exemplo, a diminuição de hiperparâmetros testados no experimento 4 e a técnica *Early Stopping* no experimento 5.

Figura 9 – Arquitetura rede neural convulcacional 03



Fonte: elaborado pelo autor (2023).

Figura 10 – Arquitetura rede neural convulciao-
nal 04



Fonte: elaborado pelo autor (2023).

6 CONCLUSÕES

A doença de Alzheimer é o tipo mais comum de demência no mundo, representando cerca de 60% a 80% dos casos (GAUGLER *et al.*, 2022). É uma das principais causas de morte entre os idosos. Mata mais do que o câncer de mama e o câncer de próstata juntos (GAUGLER *et al.*, 2022). Este trabalho teve como principal objetivo propor e avaliar uma solução para a detecção de diagnósticos da doença de Alzheimer em imagens de ressonância magnética, utilizando técnicas de aprendizagem de máquina e aprendizado profundo. Desta forma, foram implementados ao todo 15 modelos de aprendizado de máquina e 4 diferentes arquiteturas de redes neurais convolucionais.

Em relação às principais dificuldades encontradas na execução deste trabalho, pode-se destacar a dificuldade em conseguir a base de dados e a capacidade do hardware utilizado para os experimentos. Os experimentos foram dificultados pela baixa capacidade do hardware. Os experimentos poderiam ter levado em considerações mais hiperparâmetros para cada modelo.

Pode-se destacar como pontos positivos deste trabalho: a aplicação de diversas técnicas de aprendizagem de máquina e aprendizagem profunda, a contribuição com a comunidade científica na classificação da doença de Alzheimer, a aplicação dos modelos em diferentes escalas de cores e o tratamento do desbalanceamento da base de dados utilizando a SMOTE. Vale ressaltar que todas as implementações realizadas neste trabalho podem ser encontradas em um repositório público no *GitHub*.

Para trabalhos futuros, pretende-se explorar outras bases de dados mais robustas (com mais atributos), visando avaliar e melhorar o desempenho dos modelos aplicados. Pode-se também ser aplicada outras estratégias de balanceamento, considerando o problema de desbalanceamento. Além disso, pode-se estudar a otimização de mais hiperparâmetros para os modelos e aplicar outras arquiteturas de redes neurais. Outros pontos para serem estudados seria classificar a doença em todas as suas fases e recorrer às técnicas de transferência de conhecimento.

REFERÊNCIAS

- ABRAZ. **Associação Brasileira de Alzheimer**. 2022. Disponível em: <https://abraz.org.br/>. Acesso em: 12 out. 2022.
- ACADEMY, D. S. **Deep Learning Book**. 2022. Disponível em: <https://www.deeplearningbook.com.br/>. Acesso em: 12 out. 2022.
- ALTAY, F.; SÁNCHEZ, G. R.; JAMES, Y.; FARAONE, S. V.; VELIPASALAR, S.; SALEKIN, A. Preclinical stage alzheimer's disease detection using magnetic resonance image scans. In: **Proceedings of the AAAI Conference on Artificial Intelligence**. [S. l.: s. n.], 2021. v. 35, n. 17, p. 15088–15097.
- ALZ. **Ajuda sobre Alzheimer e demência | Brasil | Alzheimer's Association**. 2022. Disponível em: <https://www.alz.org.br/demencia-alzheimer-brasil.asp>. Acesso em: 12 out. 2022.
- BARROS, A. C.; LUCATELLI, J. F.; MALUF, S. W.; ANDRADE, F. M. d. Influência genética sobre a doença de alzheimer de início tardio. **Archives of Clinical Psychiatry (São Paulo)**, SciELO Brasil, v. 36, p. 16–24, 2009.
- BRAINS. **Extra-Trees: Árvores extremamente aleatórias - brains**. 2023. Disponível em: <https://brains.dev/2023/extra-trees-arvores-extremamente-aleatorias>. Acesso em: 08 jul. 2023.
- CHEN, Z.; ZHONG, C. Decoding alzheimer's disease from perturbed cerebral glucose metabolism: implications for diagnostic and therapeutic strategies. **Progress in neurobiology**, Elsevier, v. 108, p. 21–43, 2013.
- CHOLLET, F. **Deep learning with Python**. [S. l.]: Simon and Schuster, 2021.
- DATATECHNOTES. **Classification Example with Ridge Classifier in Python**. 2023. Disponível em: <https://www.datatechnotes.com/2020/07/classification-example-with-ridge-classifier-in-python.html#:~:text=The%20Ridge%20Classifier%2C%20based%20on>. Acesso em: 07 jul. 2023.
- FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. d. L. F. d. **Inteligência artificial: uma abordagem de aprendizado de máquina**. Rio de Janeiro: LTC, 2011.
- GAUGLER, J.; JAMES, B.; JOHNSON, T.; REIMER, J.; SOLIS, M.; WEUVE, J.; BUCKLEY, R. F.; HOHMAN, T. J. 2022 alzheimer's disease facts and figures. **ALZHEIMERS & DEMENTIA**, WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, v. 18, n. 4, p. 700–789, 2022.
- GÉRON, A. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems**. [S. l.]: "O'Reilly Media, Inc.", 2019.
- ISLAM, M.; TAHSIN, M. S.; ALAM, F.; HOSSAIN, S. S.; DEB, A.; KAR, S. Understanding convolutional neural network's behavior for alzheimer's disease on mri. In: **IEEE. 2022 IEEE 13th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)**. [S. l.], 2022. p. 0179–0185.
- KUMAR. **Ridge Classification Concepts & Python Examples**. 2023. Disponível em: <https://vitalflux.com/ridge-classification-concepts-python-examples>. Acesso em: 08 jul. 2023.

LAMFO. **Análise Discriminante Linear - LAMFO**. 2023. Disponível em: <https://lamfo-unb.github.io/2020/10/10/An%C3%A1lise-Discriminante-Linear>. Acesso em: 08 jul. 2023.

LIKEBUPT. **SMOTE - Azure Machine Learning**. 2023. Disponível em: <https://learn.microsoft.com/pt-br/azure/machine-learning/component-reference/smote?view=azureml-api-2>. Acesso em: 04 jul. 2023.

LUDERMIR, T. B. Inteligência artificial e aprendizado de máquina: estado atual e tendências. **Estudos Avançados**, SciELO Brasil, v. 35, p. 85–94, 2021.

LUVSANDORJ. **Meet HistGradientBoostingClassifier**. 2023. Disponível em: <https://towardsdatascience.com/meet-histgradientboostingclassifier-54a9df60d066>. Acesso em: 08 jul. 2023.

MELO, C. **XGBoost: aprenda este algoritmo de machine learning em python**. 2023. Disponível em: <https://sigmoidal.ai/xgboost-aprenda-algoritmo-de-machine-learning-em-python>. Acesso em: 07 jul. 2023.

MURUGAN, S.; VENKATESAN, C.; SUMITHRA, M.; GAO, X.-Z.; ELAKKIYA, B.; AKILA, M.; MANOHARAN, S. Demnet: a deep learning model for early diagnosis of alzheimer diseases and dementia from mr images. **IEEE Access**, IEEE, v. 9, p. 90319–90329, 2021.

NETO. **Oversampling ou Undersampling? Porque não os dois juntos?** 2023. Disponível em: <https://erasmo-aln.medium.com/oversampling-ou-undersampling-porque-n%C3%A3o-os-dois-juntos-124ecb7fbb6b>. Acesso em: 08 jul. 2023.

PETERSON, L. E. K-nearest neighbor. **Scholarpedia**, v. 4, n. 2, p. 1883, 2009.

SAGEMAKER, A. **Como o XGBoost funciona - Amazon SageMaker**. 2023. Disponível em: https://docs.aws.amazon.com/pt_br/sagemaker/latest/dg/xgboost-HowItWorks.html. Acesso em: 08 jul. 2023.

SILVA, I. R.; SILVA, G. S.; SOUZA, R. G. de; SANTOS, W. P. dos; ROBERTA, A. d. A. Model based on deep feature extraction for diagnosis of alzheimer’s disease. In: IEEE. **2019 International Joint Conference on Neural Networks (IJCNN)**. [S. l.], 2019. p. 1–7.

SILVA, J. **Uma breve introdução ao algoritmo de Machine Learning Gradient Boosting utilizando a biblioteca**. 2023. Disponível em: <https://medium.com/equal-lab/uma-breve-introdu%C3%A7%C3%A3o-ao-algoritmo-de-machine-learning-gradient-boosting-utilizando-a-biblioteca-311285783099>. Acesso em: 08 jul. 2023.

SKLEARN. **1.2. Linear and Quadratic Discriminant Analysis — scikit-learn 0.23.2 documentation**. 2023. Disponível em: https://scikit-learn.org/stable/modules/lda_qda.html. Acesso em: 08 jul. 2023.

SMITH, M. d. A. C. Doença de alzheimer. **Brazilian Journal of Psychiatry**, SciELO Brasil, v. 21, p. 03–07, 1999.

TECH, D. **Como funciona o algoritmo de Árvore de Decisão (Decision Tree)**. 2023. Disponível em: <https://didatica.tech/como-funciona-o-algoritmo-arvore-de-decisao>. Acesso em: 08 jul. 2023.

WEBB, G. I.; KEOGH, E.; MIIKKULAINEN, R. Naïve bayes. **Encyclopedia of machine learning**, v. 15, n. 1, p. 713–714, 2010.

WU, H.; LUO, J.; LU, X.; ZENG, Y. 3d transfer learning network for classification of alzheimer's disease with mri. **International Journal of Machine Learning and Cybernetics**, Springer, p. 1–15, 2022.