

# Aprendizado de Máquina para Diagnóstico e Classificação da Doença de Alzheimer

Rafael Braga Ennes  
Universidade Federal de São Paulo (UNIFESP)  
São José dos Campos, Brasil  
rafael.braga@unifesp.br

**Resumo**—A doença de Alzheimer atinge uma grande parcela da população idosa, cerca de 11,5% das pessoas nessa faixa etária, e, devido ao fato de que ainda não existe uma cura para essa doença, o diagnóstico é uma das partes mais importantes para o paciente que possui essa enfermidade, já que, quanto antes esse problema for identificado, melhor é o prolongamento de uma vida mais saudável para essa pessoa, desacelerando a progressão da doença. Dessa forma, a persistência em pesquisar técnicas mais eficazes e rápidas de diagnóstico se mostra importante. Portanto, esse projeto propôs, por meio do uso de dois conjuntos de dados, um com informações sobre o paciente e outro com imagens de ressonâncias magnéticas, utilizando modelos de machine learning para fazer previsões do estado dos cidadãos.

**Palavras Chaves**—inteligência artificial, alzheimer, predição, algoritmos

## I. INTRODUÇÃO E MOTIVAÇÃO

A doença de Alzheimer é a forma mais comum de demência degenerativa, se manifesta pela deterioração cognitiva e da memória é um transtorno neurodegenerativo progressivo, em que as atividades cotidianas são progressivamente afetadas, causando a perda de memória e da linguagem, e dependendo do grau de degeneração perde-se a condição de cuidar de si próprio. De acordo com o Ministério da Saúde, entre pessoas com 65 anos ou mais, cerca de 11,5% deles têm essa enfermidade [1], sendo que no Brasil, 1,2 milhões de cidadãos convivem com essa demência, e 100 mil novos casos são diagnosticados todo ano [2].

Essa doença não possui cura, apenas remédios que são responsáveis por melhorar a cognição e atenuar por algum tempo essa degeneração, ou pelo menos a velocidade com que ela acontece, e também não se sabe ao certo ainda a causa desse problema, existem indícios que sejam pré disposições genéticas, porém, as únicas coisas que já se sabem, é que a doença provoca uma perda das células cerebrais e que está intimamente ligado com a idade, em que, quanto mais velha a pessoa for, maiores são as chances de desenvolver a doença.

Devido a essa falta de informações concretas ainda sobre a doença, a tentativa de prevenção, a partir de um estilo de vida saudável, com a prática de exercícios físicos e uma alimentação saudável e balanceada, e o diagnóstico o mais cedo possível dessa enfermidade, comparecendo ao médicos em caso de estar com os sintomas, são as atitudes mais viáveis para a sociedade diminuir os problemas gerados por essa demência. Porém, o diagnóstico desse problema também não é fácil de ser feito, já que um dos métodos usados para isso é a entrevista médica, em que o especialista irá excluir as demais doenças por meio de exames de sangue e imagens, já que, pelo menos por enquanto, ainda não se descobriu um marcador

biológico para a doença. Sendo assim, esse processo de diagnóstico não é tão rápido e nem sempre eficaz, entretanto, atualmente avanços laboratoriais, com o auxílio de tecnologias, têm melhorado a acurácia dessa qualificação.

Sendo assim, é de extrema importância que a ciência, em conjunto com a tecnologia, continue a estudar e avançar em pesquisas que abordem temas relacionados a essa demência, já que diversas pessoas ainda sofrem com esse problema e não existem soluções eficientes para ele. Então se mostra necessário e interessante o estudo desse tema para que um diagnóstico mais rápido, medicações mais eficientes e até mesmo a cura dessa doença possam surgir, fazendo com que pessoas que sejam acometidas com essa demência tenham uma maior expectativa de vida e tenham mais esperança de uma vida melhor, e para aqueles que ainda não sofrem desse problema possam se prevenir da melhor forma possível.

A tecnologia pode ser uma grande aliada para o avanço de pesquisas sobre essa enfermidade, e na área da saúde em geral. A computação já vem sendo aplicada em diversas pesquisas na área da saúde, por exemplo, o uso de modelos de machine learning para prever de forma eficaz doenças cardíacas [3], e isso se deve ao fato de que computadores possuem um grande poder de manipular grande quantidade de dados, facilitando para os pesquisadores, não precisando analisar todos os dados em busca de achar correlações e regras dentro de um enorme conjunto de informações. O uso de inteligência artificial está sendo bastante requisitado na área da saúde, para poder fazer previsões de diagnósticos e agrupamento por características de certas doenças, aumentando a acurácia na identificação de algumas enfermidades humanas, podendo fazer isso de forma mais rápida e precisa do que alguns diagnósticos feitos por médicos.

Com isso, diversos modelos e algoritmos já são utilizados para classificação e diagnóstico de Alzheimer, como, Redes Neurais, SVM e Árvores de Decisão, já que não existem exames específicos para confirmar essa doença em um paciente, os métodos de inteligência computacional são ótimas opções para otimizar essa diagnósticos com maior acurácia.

## II. CONCEITOS FUNDAMENTAIS

Alzheimer é uma doença progressiva do cérebro que afeta a memória, o pensamento e o comportamento das pessoas. É a forma mais comum de demência em idosos e afeta milhões de pessoas em todo o mundo. Até hoje em dia, é um problema que não possui cura, apenas remédios que ajudam a melhorar a qualidade de vida das pessoas, fazendo com que essa enfermidade demore mais a progredir.

Inteligência Artificial refere-se ao avanço tecnológico que tenta desenvolver sistemas para simular a inteligência humana, encontrando padrões e regras em bases de dados, fazendo com que essas tecnologias tenham capacidade de realizar tarefas complexas, permitindo que esses problemas possam ser automatizados a fim de facilitar trabalhos que são desempenhados por seres humanos, como, aprendizado, tomada de decisões e compreensão de linguagem. Por esse motivo, essa tecnologia tem o potencial de melhorar em diversos âmbitos da sociedade, como na automatização da indústria, detecção de fraudes em transações bancárias e diagnósticos médicos.

Machine Learning, ou Aprendizado de Máquina, é uma abordagem da inteligência artificial que permite que computadores aprendam e melhorem automaticamente a partir de dados, sem serem explicitamente programados. Em vez de seguir instruções específicas, os algoritmos de Machine Learning são projetados para aprender padrões e tomar decisões com base nas informações disponíveis.

Esses algoritmos são treinados usando conjuntos de dados, nos quais são expostos a exemplos e informações relevantes para a tarefa em questão. Durante o treinamento, eles procuram padrões e relações nos dados, ajustando seus modelos e parâmetros para otimizar seu desempenho.

Existem diferentes tipos de abordagens de Machine Learning, como o aprendizado supervisionado, onde os algoritmos aprendem a partir de exemplos rotulados, o aprendizado não supervisionado, no qual os algoritmos encontram padrões em dados não rotulados, e o aprendizado por reforço, em que os algoritmos aprendem a tomar ações para maximizar uma recompensa.

Uma árvore de decisão é um algoritmo de aprendizado de máquina supervisionado que pode ser usado tanto para classificação quanto para regressão. Ela é uma representação em forma de árvore de regras de decisão que são aprendidas a partir dos dados de treinamento.

No caso da classificação, uma árvore de decisão divide o conjunto de dados em subconjuntos menores com base em diferentes atributos, buscando criar grupos homogêneos de amostras com a mesma classe. Cada nó interno da árvore representa uma decisão baseada em um atributo, e cada folha representa uma classe ou uma decisão final. Durante o treinamento, a árvore de decisão aprende quais atributos são mais informativos para tomar decisões de classificação. Isso é determinado pelo cálculo de métricas, como o ganho de informação ou a impureza de Gini, que mede a qualidade da divisão em cada nó.

Random Forest é um algoritmo de aprendizado de máquina baseado em ensemble, que combina várias árvores de decisão para tomar decisões mais precisas. Cada árvore é treinada em uma amostra aleatória dos dados de treinamento e faz uma previsão. A previsão final é determinada pela média ou votação das previsões individuais das árvores. A Random Forest é robusta, lida bem com dados de alta dimensionalidade e é capaz de lidar com problemas de classificação e regressão. Ela é amplamente utilizada em várias aplicações, como detecção de fraudes, diagnóstico médico e análise de dados.

O KNN (K-Nearest Neighbors) é um algoritmo de aprendizado de máquina supervisionado. Ele se baseia no princípio de que amostras semelhantes tendem a estar

próximas umas das outras no espaço de atributos.

No KNN, as amostras são representadas como pontos em um espaço de atributos multidimensional, onde cada atributo corresponde a uma característica das amostras. O algoritmo classifica ou prevê o valor de uma nova amostra com base nas classes ou nos valores das amostras mais próximas a ela.

SVM (Support Vector Machine) é um algoritmo de aprendizado de máquina que separa dados em grupos usando uma linha ou superfície. Ele busca encontrar a melhor separação entre os grupos, maximizando a distância entre os pontos de cada grupo. O SVM também pode lidar com dados que não podem ser separados por uma linha reta, utilizando truques matemáticos chamados de funções de kernel. É um algoritmo poderoso para classificação e regressão, com ampla aplicação em diversas áreas.

Redes neurais convolucionais (CNNs) são um tipo de algoritmo de aprendizado de máquina inspirado no funcionamento do cérebro humano. Elas são especialmente eficazes em tarefas de visão computacional, como reconhecimento de objetos em imagens. As CNNs são compostas por camadas de neurônios especializados chamados de filtros convolucionais, que extraem características importantes das imagens. Essas características são combinadas em camadas subsequentes para realizar a classificação ou detecção de objetos. As CNNs revolucionaram a área de visão computacional e têm sido amplamente utilizadas em aplicativos como carros autônomos, diagnóstico médico e reconhecimento facial.

Acurácia é uma métrica usada para avaliar o desempenho de um modelo de aprendizado de máquina em tarefas de classificação. Ela representa a porcentagem de previsões corretas feitas pelo modelo em relação ao total de amostras. Quanto maior a acurácia, melhor é o desempenho do modelo.

A função de loss (perda) é uma medida do erro do modelo durante o treinamento. Ela representa o quanto as previsões do modelo estão distantes dos valores reais dos dados de treinamento. O objetivo é minimizar a loss durante o treinamento para que o modelo possa fazer previsões mais precisas. Diferentes algoritmos e problemas têm diferentes tipos de função de loss, como a mean squared error (MSE) para regressão ou a cross-entropy para classificação.

Grid search é uma técnica usada no processo de ajuste de hiperparâmetros em algoritmos de aprendizado de máquina. Ele consiste em definir um conjunto de valores para cada hiperparâmetro do modelo e avaliar o desempenho do modelo para todas as combinações possíveis desses valores. Isso permite encontrar a combinação de hiperparâmetros que produz o melhor desempenho em termos de uma métrica de avaliação, como acurácia ou F1-score.

### III. TRABALHOS RELACIONADOS

Diversos trabalhos já propuseram fazer o diagnóstico de doenças a partir do uso de técnicas de machine learning (ML). A pesquisa feita por Parthiaban e Srivatsa [4] tem por objetivo o diagnóstico de doenças cardíacas em pacientes diabéticos com o uso de métodos de ML, como, Naive Bayes e SVM usando o pacote de software WEKA, e foi usado um conjunto de dados supervisionado em que 142 dos 500 pacientes tinham a doença, e nesse caso o SVM se performou melhor que o Naive Bayes. Fathima e Manimeglai [5] propuseram prever a

doença da dengue também utilizando para a mineração dos dados o SVM, utilizando um conjunto de dados de 29 atributos e 5000 amostras.

No [6] os autores tiveram como objetivo diagnosticar pessoas com problemas na tireóide, e para solucionar esse problema eles usaram arquiteturas de redes neurais artificiais, o conjunto de dados supervisionado possui 19 categorias clínicas para a enfermidade, e as redes neurais estudadas foram as multilayer perceptron (MLP) com o uso de back-propagation. O desenvolvimento das redes neurais possibilitaram uma empolgação e uma grande ampliação dos estudos na área da biomedicina, como é abordado em [7], principalmente com o ampliação do conhecimento de redes neurais convolucionais, permitindo, também, fazer previsões a partir de imagens.

E também já existem vários estudos de doenças usando análise de imagens com machine learning, como foi feito nos casos [8][9], em que houve aplicações com sucesso de técnicas de machine learning convencionais para imagens médicas, em que foi usado inteligência artificial para reconhecimento de padrões em radiografias, para a identificação de um possível câncer.

Especificamente, para a área de estudos neurológicos, a partir da análise de imagens de ressonância magnética, [10] e [11] aplicaram machine learning para estudar sobre o cérebro de diversos aspectos, principalmente para a identificação de problemas psiquiátricos.

Para a doença de Alzheimer também existem pesquisas usando inteligência artificial para o diagnóstico dessa enfermidade, Neelaveni e Devasana [12] usaram SVM e árvore de decisão para encontrar os resultados e comparar os métodos. E o [13] utilizou de um conjunto de dados com 12 atributos e sem imagens, um conjunto supervisionado, e o grupo utilizou também de diversos métodos para preparação dos dados e depois para a análise deles, predizendo entre dois grupos de pacientes, os dementes e os não dementes.

Já a pesquisa [14], utiliza imagens de ressonâncias magnéticas para fazer a predição entre três classes de pacientes usando modelos de redes neurais convolucionais, e é apresentado no trabalho as diferenças de resultados ao se modificar certos parâmetros.

#### IV. OBJETIVOS

Esse artigo tem por objetivo, usando dois conjuntos de dados supervisionado, um que possui imagens de ressonância magnética de possíveis pacientes com a doença de Alzheimer, e um outro conjunto de dados, com o diagnóstico e alguns atributos que descrevem os pacientes, fazer o uso de diferentes modelos de aprendizado de máquinas, para tentar encontrar a melhor solução entre elas, em cada um dos conjuntos e fazer uma análise, dentre os atributos do segundo conjunto de dados, para ver aqueles que têm maior relação com o estado da pessoa. Essa pesquisa se mostra muito importante, já que obtendo resultados positivos, em que a acurácia dos modelos de predição sejam altos, poderia ser aplicado em casos reais para um diagnóstico precoce dessa enfermidade, havendo possibilidade de tratamentos que amenizem os efeitos da doença, aumentando a expectativa e a qualidade de vida da pessoa.

#### V. MÉTODOS EXPERIMENTAIS

Para que o projeto se concretize serão usadas bibliotecas de processamento de dados e imagens (Numpy, Pandas, Scikit-learn e TensorFlow), também serão testados diferentes modelos de machine learning, como, SVM, Árvore de Decisão e Redes Neurais.

Os passos que devem ser realizados durante todo o processo da construção do experimento e artigo são: pré-processamento dos dados (fazendo as mudanças necessárias para cada modelo que será utilizado), extração de padrões (utilizando os diferentes modelos para cada conjunto), pós-processamento (análise dos resultados obtidos e busca pelos melhores modelos). A partir disso será feita a conclusão dos melhores métodos e técnicas a serem utilizadas, e também a análise, se é importante o prosseguimento de pesquisas nesta área, caso os resultados sejam bons. Abaixo está o diagrama de bloco com os passos seguidos para o experimento (Fig.1).

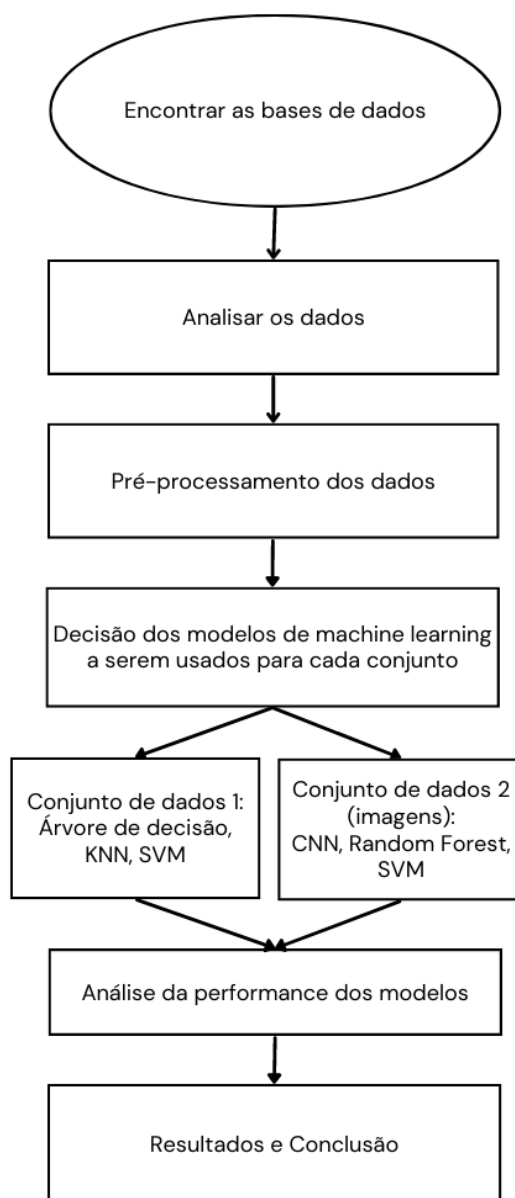


Fig. 1. Diagrama de bloco.

### A. Banco de dados

Como dito anteriormente, para esse experimento foram utilizados dois diferentes bancos de dados, o primeiro com dados sobre os pacientes, como, idade, anos de educação e status socioeconômico, que buscam juntar informações sobre as pessoas para que seja possível fazer a predição sobre a doença. Já o segundo conjunto de dados é composto por imagens de ressonância magnética dos pacientes, ambos os conjuntos de dados são rotulados com as devidas classificações dos possíveis enfermos.

A primeira base de dados possui 373 instâncias e 10 atributos para cada paciente (Tabela 1), e a classificação final dos grupos se divide entre três grupos, Demented, Nondemented e Converted (pessoas que haviam sido classificadas anteriormente como saudáveis e em outro exame como doente). Dentro desse conjunto de dados existiam 21 dados faltantes, 19 para o atributo SES (Socioeconomic Status) e 2 para o MMSE (Mini-Mental State Examination), que foram substituídos por valores que faziam mais sentido de

acordo com a distribuição dos dados de cada um desses atributos, que foi a substituição pela mediana e pela média, respectivamente.

O segundo conjunto é composto por 6400 imagens de ressonâncias magnéticas, em que os dados já estão divididos pelo nível de gravidade de Alzheimer: NonDemented, ModerateDemented, MildDemented e VeryMildDemented (Fig. 1). Para o pré processamento desses dados foi necessário formatar as imagens (Fig. 2), para ajustar elas e deixá-las no formato ideal para aplicar os algoritmos de machine learning, com isso, foi feito uma recoloração das imagens, as transformando em escalas de cinza, e também foi feito o redimensionamento delas. E também deve ser notado que esse segundo conjunto possui dados desbalanceados, já que, principalmente o caso dos pacientes que estão com o nível de Alzheimer mais grave é bem menor que os demais, e esse fato será importante para o cálculo das métricas de avaliação dos modelos.

Atributos	Atributos: Consiste em 10 atributos que são descritos da seguinte forma:
	Descrição
Group	Nível de demência do paciente: Converted, Demented, Nondemented
M/F	Gênero
Age	Idade ( Em anos)
EDUC	Anos de educação
SES	Status socioeconômico, classificado de 1 (maior status) a 5 (menor status)
MMSE	Valor da examinação do estado minimental (0 = pior caso e 30 = melhor caso)
CDR	Avaliação clínica da demência (0 = não é demente, 0.5 = demência muito leve, 1 = demência leve, 2 = demência moderada)
eTIV	Volume estimado intracranial, mm3
nWBV	Volume normalizado do cérebro inteiro
ASF	Fator de escala do Atlas

Tabela. 1. Identificação e explicação sobre os 10 atributos presentes no primeiro conjunto de dados.

### B. Protocolo Experimental

Foram usados cinco técnicas diferentes de ML para fazer as predições dos modelos: Árvore de Decisão, K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Random Forest e Convolutional Neural Network (CNN). Sendo que Árvore de Decisão, KNN e SVM foram usados para o primeiro conjunto de dados, e o Random Forest, SVM e CNN foram escolhidos para fazer a análise do conjunto de imagens.

Para que as implementações tivessem resultados melhores, em todos os modelos produzidos para a predição do primeiro conjunto, foi usada a técnica do Grid Search em que diferentes parâmetros foram testados para encontrar os que geraram mais acertos, maximizando o desempenho do modelo.

Como medidas de avaliação dos modelos implementados foi utilizado a taxa de acurácia, que representa a porcentagem de previsões corretas feitas pelo modelo em relação ao total de amostras, a precisão, que é a proporção de previsões corretas positivas em relação ao total de previsões positivas, o recall, que representa a proporção de previsões corretas positivas em relação ao total de amostras positivas reais e a matriz de

confusão, que é uma tabela que mostra a contagem de verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos do modelo.

E para finalizar a análise dos resultados foi utilizado, como um teste de diferença estatística, o método ANOVA (Análise de Variância) que é utilizado para determinar se há diferença estatisticamente significativa entre as médias de três ou mais grupos/métodos. Se a variação entre os grupos for significativamente maior do que a variação dentro dos grupos, isso indica que existe uma diferença significativa entre pelo menos um par de modelos. Ele é uma ferramenta importante na análise de comparações múltiplas em experimentos e estudos que envolvem diferentes tratamentos ou abordagens.

### C. Protocolo de Validação

O principal objetivo de um algoritmo de aprendizado de máquina é ser capaz de generalizar dados novos, pois diferentemente dos dados de treinamento, nós não sabemos os valores dos rótulos deles, por isso é necessário colocar uma das partes do conjunto de dados para o teste do modelo utilizado.

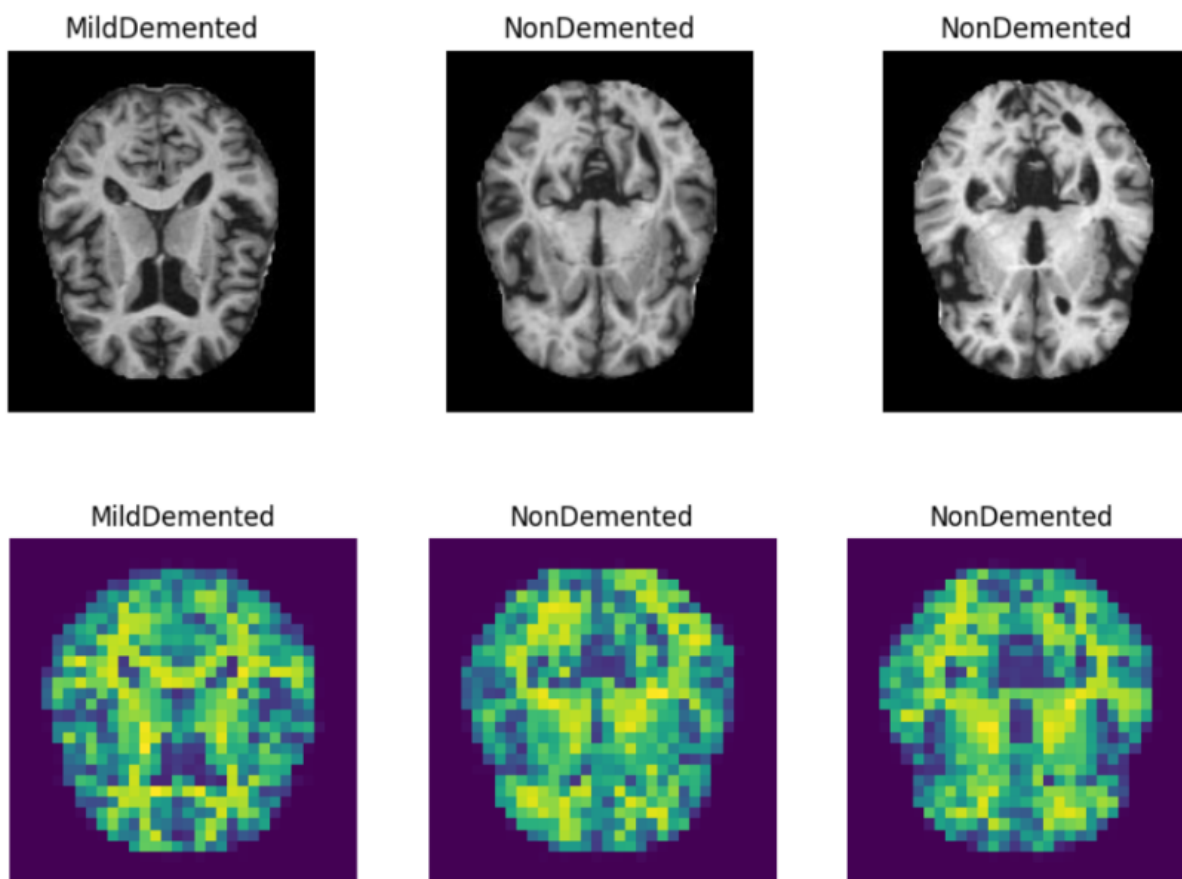


Fig. 2 e Fig.3. O primeiro conjunto de imagens mostra diferentes ressonâncias magnéticas com os seus respectivos rótulos, antes do pré processamento. Já o segundo conjunto apresenta as mesmas imagens, porém após o redimensionamento e recoloração delas

Um dos métodos para tentar melhorar a predição de um modelo é o K-fold cross validation, em que o modelo é treinado K vezes (número que é escolhido anteriormente) em K-1 subsets dos dados originais e é testado no subset restante em cada vez, para cada uma das abordagens o conjunto de teste muda. Assim, se obtêm K modelos de qualidades a serem analisados, um de cada abordagem, e na maioria das vezes é feito uma média desses resultados para que se obtenha um resultado mais perto do geral.

O cross-validation ajuda a evitar o overfitting, pois avalia o desempenho do modelo em dados não vistos durante o treinamento. Ele fornece uma estimativa mais realista do desempenho do modelo em cenários de uso real. Além disso, o cross-validation também pode ser usado para comparar diferentes modelos ou ajustar hiperparâmetros, ajudando a escolher a melhor configuração para o problema em questão.

Para todos os modelos testados para o primeiro conjunto de dados deste trabalho (dados dos pacientes) foi usado o método de validação K-fold cross validation, em que o K foi definido como 5, ou seja, foram produzidos cinco folds a cada iteração em que apenas um era para teste, e nas iterações seguintes era alternado o fold que deveria ser usado para validar o modelo.

## VI. RESULTADOS E DISCUSSÃO

Para análise dos resultados foram usadas algumas medidas de avaliação que serão apresentados os valores finais de cada

um deles em seguida.

Pode ser notado que, juntamente com o uso do Grid Search, para encontrar os melhores hiperparâmetros para cada caso, foi possível obter bons resultados para a predição que era desejada. Todos os modelos testados obtiveram resultados de suas acurácias maiores que 80%, o que mostra que nenhum deles desempenhou de maneira ruim.

Como resultado do Grid Search usado nos modelos do primeiro conjunto de dados obtivemos como os melhores hiperparâmetros a serem usados nessa base de dados (Tabela 2):

Modelo	Melhores hiperparâmetros obtidos no Grid Search
	<i>Hiperparâmetros</i>
Árvore de Decisão	'criterion': 'gini', 'max_depth': 8, 'min_samples_leaf': 1, 'min_samples_split': 4, 'splitter': 'random'
KNN	'knn_n_neighbors': 1
SVM	'C': 10, 'gamma': 1, 'kernel': 'rbf'

Tabela. 2. Resultados do Grid Search com os melhores hiperparâmetros para essa base de dados.

Devido ao fato de que foram testados diferentes casos em

cada modelo, utilizando diferentes parâmetros, a técnica de Grid Search foi capaz de encontrar os hiperparâmetros que mais maximizavam os resultados, dessa forma, foi possível alcançar bons resultados.

Além da técnica de Grid Search foi utilizado o k-fold cross validation, em que o conjunto de dados é dividido em k folds e

1 deles é utilizado para teste e o restante para treino, e esse experimento é feito k vezes até que todos os dados tenham sido usados como teste uma vez, e podemos ver na tabela (Tabela 3) os resultados de cada um desses testes, com a sua médias e desvio padrão.

Modelo	Acurácias obtidas a partir do k-fold cross validation (Primeiro conjunto de dados)						
	Acurácia 1	Acurácia 2	Acurácia 3	Acurácia 4	Acurácia 5	Média	Desvio Padrão
Árvore de Decisão	84.38	86.15	90.90	86.67	85.25	86.67	0.0226
KNN	88.24	86.27	90.00	94.00	90.00	89.70	0.0255
SVM	86.27	90.20	90.00	94.00	90.00	90.09	0.0244

Tabela. 3. Resultados do k-fold cross validation.

Foram performados os modelos de machine learning propostos, com o objetivo de fazer a predição de pacientes com Alzheimer, foi produzido uma tabela com os resultados obtidos de acurácia de cada técnica utilizada (Tabela 4 e Tabela 5).

Modelo	Acurácias obtidas a partir do melhor resultado de cada técnica (Primeiro conjunto de dados)
	Acurácia
Árvore de Decisão	86.67
KNN	89.70
SVM	90.09

Tabela. 4. Resultados da melhor acurácia obtida por método para o primeiro conjunto.

Após obter os resultados das performances de cada um dos métodos podemos ver que o modelo que utilizou o SVM foi o que obteve melhores valores de acurácia média no k-fold cross validation, o que já era esperado pois é um método que performa bem em grande parte dos conjuntos de dados, no entanto as outras técnicas também tiveram resultados próximos, o que mostra que também foram bem em suas predições.

E com esses dados obtidos foi possível utilizar da técnica ANOVA, que é um teste de variância entre os resultados de cada modelo para verificar se há um grande diferença entre os métodos de machine learning utilizados ou não, tomando como base os cinco resultados obtidos no k-fold cross validation de cada modelo. E com isso, foi obtida a resposta de que não há uma diferença significativa entre os métodos usados.

Podemos verificar que para as predições feitas do conjunto de dados das imagens de ressonância magnética também alcançaram bons resultados, sendo o maior deles o experimento com CNN, em que a acurácia foi de 91,51%, e isso já era esperado, pois o CNN é um modelo que performa

muito bem geralmente na análise e reconhecimento de padrões de imagens. Entretanto, todos os métodos obtiveram resultados acima de 80%, o que também são resultados bons.

Modelo	Acurácias obtidas a partir do melhor resultado de cada técnica (Segundo conjunto de dados)
	Acurácia
CNN	92.83
Random Forest	92.52
SVM	82.71

Tabela. 5. Resultados da melhor acurácia obtida por método para o segundo conjunto.

Porém, pelo fato de que estamos lidando com uma base de dados desbalanceada, a acurácia não performa muito bem, sendo assim, também foi usada a métrica F1-score (Tabela 6), que é uma média harmônica entre o precision e o recall, que em geral, lida melhor com dados que possuem classes com números muito diferentes de instâncias.

Pode ser observado que novamente o algoritmo da CNN foi a que performou melhor para a métrica F1-score, usada para essa base de dados.

Modelo	Valores obtidos usando o F1-score
	F1-score
CNN	0.876
Random Forest	0.819
SVM	0.864

Tabela. 6. Resultados da métrica F1-score sendo usada para os dados desbalanceados.

## VII. CONCLUSÕES E TRABALHOS FUTUROS

Tendo em vista a importância do diagnóstico o mais cedo possível em caso de doenças como o Alzheimer, para que se possa obter resultados melhores na postergação dos problemas causados por essa enfermidade, é necessário, cada vez mais, encontrar métodos que facilitem e deixem menos custoso, em tempo e preço, para o paciente conseguir esse diagnóstico.

Neste trabalho foi proposto o uso de machine learning para encontrar modelos que permitissem fazer uma boa predição da situação dos pacientes, a partir de dois conjuntos de dados, um com as informações sobre pacientes e outro com imagens de ressonâncias magnéticas.

Foi possível alcançar bons resultados em ambos conjuntos de dados e em todas as técnicas, sendo utilizados modelos e métricas que faziam sentido de acordo com o caso. Apenas deve ser ressaltado o fato de que haviam poucas instâncias no conjunto de dados 1 (apenas 373), o que pode ter sido um entrave para que a acurácia não fosse ainda maior, já que não teve muitos dados para treinamento.

Porém, de forma geral, pode se dizer que o experimento cumpriu com os seus objetivos iniciais.

Para trabalhos futuros é almejado que seja obtido base de dados maiores para que os modelos possam obter resultados melhores, além de utilizar algoritmos de predições e métricas de avaliações diferentes para poder explorar outros problemas, por exemplo o uso de outros métodos ensembles, que é um método que utiliza de diferentes modelos de aprendizado de máquina em conjunto, com objetivo de obter resultados preditivos melhores do que poderia ser alcançado por um método sozinho. Ajudando assim as áreas de pesquisas e da medicina com o avanço de conhecimentos nessa área, e também tentar ajudar de alguma forma as pessoas que sofrem com o mal de Alzheimer.

## REFERENCES

- [1] Ministério da Saúde. (2021). Conhecer a demência, conhecer o Alzheimer: o poder do conhecimento – Setembro, Mês Mundial do Alzheimer [Online]. Disponível em: <https://bvsmis.saude.gov.br/conhecer-a-demencia-conhecer-o-alzheimer-o-poder-do-conhecimento-setembro-mes-mundial-do-alzheimer/>
- [2] Fio Cruz. (2019). Alzheimer acomete 11,5% da população idosa do País [Online]. Disponível em: <https://www.canalsaude.fiocruz.br/noticias/noticiaAberta/alzheimer-acomete-115-da-populacao-idosa-do-pais>
- [3] S. Mohan, C. Thirumalai e G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access, vol. 7, pp. 81542-81554, 2019.
- [4] Parthiban, G. and Srivatsa, S.K. (2012) Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients. International Journal of Applied Information Systems (IJ AIS), 3, 25-30.
- [5] Fathima, A.S. and Manimeglai, D. (2012) Predictive Analysis for the Arbovirus-Dengue using SVM Classification. International Journal of Engineering and Technology, 2, 521-527.
- [6] Sharpe PK, Solberg HE, Rootwelt K, Yearworth M. Artificial neural networks in diagnosis of thyroid function from in vitro laboratory tests. Clin Chem. 1993 Nov;39(11 Pt 1):2248-53. PMID: 8222217.
- [7] Bishop CM. 1995. Neural Networks for Pattern Recognition. New York: Oxford Univ. Press
- [8] Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.H.; Aerts, H.J.W.L. Artificial intelligence in radiology. Nat. Reviews. Cancer 2018, 18, 500–510.
- [9] Reig, B.; Heacock, L.; Geras, K.J.; Moy, L. Machine learning in breast MRI. J. Magn. Reson. Imaging JMRI 2020, 52, 998–1018.
- [10] Eslami, T.; Almuqhim, F.; Raiker, J.S.; Saeed, F. Machine learning methods for diagnosing autism spectrum disorder and attention-deficit/hyperactivity disorder using functional and structural MRI: A survey. Front. Neuroinformatics 2021, 14, 575999.
- [11] Zhang, Z.; Li, G.; Xu, Y.; Tang, X. Application of artificial intelligence in the MRI classification task of human brain neurological and psychiatric diseases: A scoping review. Diagnostics 2021, 11, 1402.
- [12] J. Neelaveni and M. S. G. Devasana, "Alzheimer Disease Prediction

using Machine Learning Algorithms," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, 2020, pp. 101-104, doi: 10.1109/ICACCS48705.2020.9074248.

- [13] Kavitha C, Mani V, Srividhya SR, Khalaf OI, Tavera Romero CA. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. Front Public Health. 2022 Mar 3;10:853294. doi: 10.3389/fpubh.2022.853294. PMID: 35309200; PMCID: PMC8927715.

- [14] A. W. Salehi, P. Baglat, B. B. Sharma, G. Gupta and A. Upadhyay, "A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI," 2020 International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2020, pp. 156-161, doi: 10.1109/ICOSEC49089.2020.9215402.