

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE BAURU

ANDRÉ LUIZ DA SILVA JUNIOR

ABORDAGEM DE APRENDIZADO PROFUNDO PARA CLASSIFICAÇÃO DE PARKINSON
POR MEIO DE SINAIS DE VOZ

Unesp Bauru
2023

André Luiz da Silva Junior

**ABORDAGEM DE APRENDIZADO PROFUNDO PARA CLASSIFICAÇÃO DE PARKINSON
POR MEIO DE SINAIS DE VOZ**

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Bacharel em Ciências da Computação do Campus de Bauru, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Bacharel em Ciências da Computação .

Orientador: Profº Dr. Clayton Reginaldo Pereira

Unesp Bauru
2023

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE BAURU

ANDRÉ LUIZ DA SILVA JUNIOR

ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO PARTE DO
REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE "GRADUANDO EM BACHAREL
EM CIÊNCIAS DA COMPUTAÇÃO "

APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE GRADUAÇÃO EM
BACHAREL EM CIÊNCIAS DA COMPUTAÇÃO

Profº Dra. ANDRÉA CARLA GONÇALVES VIANNA
Coordenador

BANCA EXAMINADORA:

Profº Dr. Clayton Reginaldo Pereira
Orientador/UNESP-Bauru

Profº Dra. Simone das Graças Domingues Prado
UNESP-Bauru

Profº Dr. Douglas Rodrigues
Membro Externo

Novembro , 2023

RESUMO

A doença de Parkinson é um distúrbio neurológico progressivo que afeta o cérebro. O desenvolvimento de sistemas especializados capazes de diagnosticar automaticamente e com alta precisão estágios iniciais da doença de Parkinson com base em sinais de fala representaria uma contribuição importante para o setor de saúde. Para isso, foram analisados modelos de *machining learning* como a regressão logística, *support vector machine*, *K-Neighbors*, *Random Forest*, *Decion Tree*, *Naive Bayes* e *XGBoost*, além disso foi utilizado técnicas de *deep learning* como a Rede Neural Convolucional e a Rede Neural Recorrente. Essas abordagens foram aplicadas a três conjuntos de dados com características próprias para um mais amplo estudo dos casos. Dois desses conjuntos de dados foram analisados com o uso de *machining learning*, e um com a utilização de *deep learning* com séries temporais, após o treino o modelo com o melhor desempenho foi selecionado. Os resultados revelaram que o modelo de teve um desempenho melhor nos dois conjunto de dados que utilizaram *machining learning* foi a Regressão Linear, e no *deep learning* foi a Rede Neural Convolucional. Os resultados obtidos revelaram uma acurácia de classificação de 92.31% e 86.84% para os dois conjuntos de dados, o que evidencia a capacidade de obter resultados excelentes mesmo com um conjunto limitado de dados. Além disso, foram observadas precisões de 91.43% e 86.72%, *recalls* de 100.00% e 97.37%, e *F1-Scores* de 95.52% e 91.74%. Esses valores corroboram a qualidade dos resultados alcançados. Por sua vez, a Rede Neural Convolucional apresentou uma acurácia de classificação de 86.69%, precisão de 86.83%, *recall* de 88.20% e *F1-Score* de 87.51%, mesmo com um número reduzido de épocas de treinamento. Isso evidencia a eficácia desse modelo como uma alternativa valiosa para a classificação. Esses resultados indicam que as classificações terão um papel significativo no avanço do setor médico. É importante ressaltar a relevância de uma análise cuidadosa na escolha do modelo a ser utilizado, pois diferentes conjuntos de dados podem apresentar resultados mais favoráveis em modelos distintos.

PALAVRAS-CHAVE: Doença de Parkinson, Diagnóstico , Machining learning, Deep learning, Rede Neural Convolucional, Rede Neural Recorrente.

LISTA DE ILUSTRAÇÕES

Figura 1	Redes Neurais em Camadas	19
Figura 2	Redes Neurais Recorrentes	21
Figura 3	Exemplo de <i>Support Vector Machine</i>	25
Figura 4	Distribuição do Primeiro <i>Dataset</i>	35
Figura 5	Primeira Distribuição de <i>Features</i> do Primeiro <i>Dataset</i>	36
Figura 6	Segunda Distribuição de <i>Features</i> do Primeiro <i>Dataset</i>	36
Figura 7	Análise T-SNE do Primeiro <i>Dataset</i>	37
Figura 8	Distribuição do Segundo <i>Dataset</i>	37
Figura 9	Distribuição de Gênero Segundo <i>Dataset</i>	38
Figura 10	Distribuição do Número de Pulsações do Segundo <i>Dataset</i>	38
Figura 11	Distribuição do Número de Pulsações médias do Segundo <i>Dataset</i>	39
Figura 12	Distribuição do Terceiro <i>Dataset</i>	39
Figura 13	Gráfico de Correlação do Primeiro <i>Dataset</i>	40
Figura 14	<i>Features</i> com mais Correlações do Primeiro <i>Dataset</i>	41
Figura 15	Vâriança dos Componentes de PCA do Segundo <i>Dataset</i>	41
Figura 16	<i>Scatterplot</i> do PCA dos dois principais <i>features</i> selecionados do Segundo <i>Dataset</i>	42
Figura 17	Dado Temporais	43
Figura 18	Transformações Fourier	43
Figura 19	Coeficientes <i>Filter Bank</i>	43
Figura 20	<i>Mel Frequency Cepstral Coefficients</i>	43
Figura 21	Tabela criada	44
Figura 22	Regressão Linear no Primeiro <i>Dataset</i>	46
Figura 23	Regressão Linear no Segundo <i>Dataset</i>	47
Figura 24	<i>Support Vector Machine</i> no Primeiro <i>Dataset</i>	47
Figura 25	<i>Support Vector Machine</i> no Segundo <i>Dataset</i>	48
Figura 26	K-Neighbors no Primeiro <i>Dataset</i>	48
Figura 27	K-Neighbors no Segundo <i>Dataset</i>	49
Figura 28	<i>Random Forest</i> no Primeiro <i>Dataset</i>	50
Figura 29	<i>Random Forest</i> no Segundo <i>Dataset</i>	50
Figura 30	<i>Decision Tree</i> no Primeiro <i>Dataset</i>	51
Figura 31	<i>Decision Tree</i> no Segundo <i>Dataset</i>	51
Figura 32	<i>Naive Bayes</i> no Primeiro <i>Dataset</i>	52
Figura 33	<i>Naive Bayes</i> no Segundo <i>Dataset</i>	53
Figura 34	<i>XGBoost</i> no Primeiro <i>Dataset</i>	53
Figura 35	<i>XGBoost</i> no Segundo <i>Dataset</i>	54
Figura 36	Treinamento CNN	54
Figura 37	Acurácia do Treinamento CNN	55

Figura 38 Perda do Treinamento CNN 55

Figura 39 Matriz de Confusão CNN 55

Figura 40 Treinamento RNN 56

Figura 41 Acurácia do Treinamento RNN 56

Figura 42 Perda do Treinamento RNN 56

Figura 43 Matriz de Confusão RNN 57

LISTA DE ABREVIATURAS E SIGLAS

TCC	Trabalho de Conclusão de Curso
UNESP	Universidade Estadual Paulista
IA	Inteligência Artificial
AI	Artificial Intelligence
CNNs	Convolutional Neural Networks
RNNs	Recurrent Neural Network
RNRs	Redes Neurais Recorrentes
LSTMs	Long Short Term Memory
GRUs	Gated Recurrent Unit
BERT	Bidirectional Encoder Representations
GPT	Generative Pre-trained Transformer
PLN	Processamento de Linguagem Natural
SVMs	Support Vector Machines
k-NNs	k-Nearest Neighbors
IBM	International Business Machines Corporation
DP	Doença de Parkinson
SGD	Stochastic Gradient Descent
F1	First Formant Frequency
F2	Second Formant Frequency
DDF1	Second Derivative of F1
DDF2	Second Derivative of F2
BBE	Band Bark Energies
MFCC	Mel Frequency Cepstrum Coefficient
MDVR-KCL	Mobile Device Voice Recordings at King's College London
VGG19	Visual Geometry Group 19

TWQT	Tunable Q-Factor Wavelet Transform
t-SNE	t-Distributed Stochastic Neighbor Embedding
PCA	Principal Component Analysis
TP	True Positive
TN	True Negative
FP	False Positive
FN	False Negative
FFT	Fast Fourier Transform
FBANK	Filter Bank

SUMÁRIO

1	INTRODUÇÃO	11
1.1	PROBLEMA	11
1.2	JUSTIFICATIVA	12
1.3	OBJETIVOS	12
1.3.1	OBJETIVOS GERAIS	12
1.3.2	OBJETIVOS ESPECÍFICOS	12
1.4	METODOLOGIA DE PESQUISA	13
2	REVISÃO TEÓRICA	14
2.1	Doença de Parkinson	14
2.1.1	Introdução	14
2.1.2	Sintomas	14
2.1.3	Detecção	15
2.1.4	Coleta de Dados	15
2.1.5	Necessidade de métodos de IA para detecção de Parkinson	16
2.2	Machine Learning	18
2.2.1	Redes Neurais	18
2.2.1.1	Redes Neurais Multicamada	18
2.2.1.2	Redes Neurais Convolucionais	20
2.2.1.3	Redes Neurais Recorrentes	21
2.2.2	Deep Learning	22
2.2.3	Attention	23
2.2.4	Transformers	24
2.2.5	Support Vector Machines	24
2.2.6	Decision Tree	25
2.2.7	Random Forest	26
2.2.8	Regressão Logística	26
2.2.9	XGBoost	26
2.2.10	Naive Bayes	27
2.2.11	K-Neighbors	27
3	REVISÃO BIBLIOGRAFICA	29
3.0.1	Inteligência Artificial e Medicina	29
3.0.2	A new approach: information gain algorithm-based k-nearest neighbors hybrid diagnostic system for Parkinson's disease	29
3.0.3	Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset	30

3.0.4	A generic optimization and learning framework for Parkinson disease via speech and handwritten records	31
3.0.5	A Speech-Based Hybrid Decision Support System for Early Detection of Parkinson's Disease	31
3.0.6	Early detection of Parkinson's disease using machine learning	32
4	DATASETS	33
4.0.1	Parkinson's Disease Data Set	33
4.0.2	Parkinson's Disease (PD) classification	34
4.0.3	Mobile Device Voice Recordings at King's College London (MDVR-KCL) from both early and advanced Parkinson's disease patients and healthy controls . .	34
4.1	Analise dos dados	34
4.1.1	Primeiro Dataset	35
4.1.2	Segundo Dataset	37
4.1.3	Terceiro Dataset	39
4.2	Seleção de features	40
4.2.1	Primeiro Dataset	40
4.2.2	Segundo Dataset	41
4.2.3	Terceiro Dataset	43
5	TREINAMENTO	44
5.1	Métricas utilizadas	44
5.2	Sampling dos dados	45
5.3	Treinamento dos modelos	46
5.3.1	Regressão Logística	46
5.3.1.1	Primeiro Dataset	46
5.3.1.2	Segundo Dataset	46
5.3.2	Support Vector Machine	47
5.3.2.1	Primeiro Dataset	47
5.3.2.2	Segundo Dataset	47
5.3.3	K-Neighbors	48
5.3.3.1	Primeiro Dataset	48
5.3.3.2	Segundo Dataset	49
5.3.4	Random Forest	49
5.3.4.1	Primeiro Dataset	49
5.3.4.2	Segundo Dataset	50
5.3.5	Decision Tree	50
5.3.5.1	Primeiro Dataset	50
5.3.5.2	Segundo Dataset	51
5.3.6	Naive Bayes	52
5.3.6.1	Primeiro Dataset	52

5.3.6.2	Segundo Dataset	52
5.3.7	XGBoost	53
5.3.7.1	Primeiro Dataset	53
5.3.7.2	Segundo Dataset	53
5.3.8	Deep Learning no Terceiro Dataset	54
5.3.8.1	CNN	54
5.3.8.2	RNN	55
5.4	Resultados	57
6	CONCLUSÃO	59
	REFERÊNCIAS	60

1 INTRODUÇÃO

A doença de Parkinson é uma patologia neurológica crônica e progressiva que afeta a capacidade do cérebro de controlar os movimentos corporais. Os sintomas incluem tremores, rigidez muscular, lentidão dos movimentos e instabilidade postural, o que pode resultar em dificuldades significativas na realização das atividades cotidianas.(SAÚDE, 2019)

Na área medica a inteligência artificial é usada com computadores para analisar grande volumes de dados e usando sistemas especialistas para elaborar soluções para problemas médicos, computadores tem a capacidade de armazenar e recuperar imagens de lesões já analisadas e proporciona diagnósticos mais rápidos.(LOBO, 2017)

Recentemente, o uso de técnicas de *deep learning* tem sido usadas como uma abordagem promissora para ajudar na detecção e diagnóstico do Parkinson. Em particular, o uso de sinais de voz para a classificação do Parkinson tem sido objeto de estudo, devido à sua facilidade de aquisição e potencial para ser usado em um ambiente clínico.

A doença de Parkinson causa uma vibração anormal nas cordas vocais, com a presença de ruído ou outras perturbações. Medidas são usadas pra monitorar essas disfunções, incluindo ruídos ou perturbações de frequências e amplitudes, podendo avaliar a gravidade da doença.(WANG YAN FU, 2022b)

Este trabalho tem como objetivo aplicar técnicas de *deep learning* para a classificação do Parkinson a partir de sinais de voz. A análise levará em conta dados de sinais de voz de pacientes com Parkinson e indivíduos saudáveis, com o objetivo de identificar padrões que possam ajudar no diagnóstico precoce da doença e poder diferenciar os indivíduos.

1.1 PROBLEMA

A doença de Parkinson é uma doença neurológica sem cura e progressiva que afeta principalmente o sistema motor, causando tremores, rigidez muscular, entre outros problemas. Estima-se que a doença afete entre 1% e 2% das pessoas com mais de 60 anos de idade em todo o mundo.(JANKOVIC, 2008)

O diagnóstico da doença de Parkinson é baseado principalmente na avaliação, incluindo a observação dos sintomas motores característicos da doença, assim como a realização de exames neurológicos e testes com imagens. Porém, a doença de Parkinson pode ser difícil de diagnosticar em seus estágios iniciais, já que os sintomas podem ser sutis e semelhantes a outras condições neurológicas.

A detecção precoce da doença de Parkinson é importante para a administração adequada da terapia e melhoria da qualidade de vida do paciente. Tendo isso em mente, desenvolver o uso de sinais de voz como indentificador da doença de Parkinson pode ser de grande importancia para a medicina.

Um dos desafios na utilização de sinais de voz para o diagnóstico do Parkinson é a variação natural entre indivíduos saudáveis e pacientes com Parkinson. Além disso, os sintomas da doença podem variar de um paciente para outro, o que pode afetar a precisão da classificação baseada em sinais de voz se o paciente não apresentar esse sintoma na garganta. Outra coisa que pode comprometer são ruídos externos que podem afetar a pesquisa, tudo isso deve ser levado em consideração.

1.2 JUSTIFICATIVA

A doença de Parkinson é uma doença sem cura e progressiva que afeta milhões de pessoas em todo o mundo. O diagnóstico rápido da doença é fundamental para uma terapia adequada e melhoria da qualidade de vida do paciente. No entanto, a detecção precoce da doença de Parkinson ainda é um desafio para os médicos e pesquisadores por seus sinais muito sutis em seu início.

Assim, a utilização de sinais de voz como detector da doença de Parkinson tem sido cada vez mais objeto de estudo em várias pesquisas. A fala é uma das principais funções motoras afetadas pela doença de Parkinson, e a análise de sinais de voz pode fornecer informações sobre a condição do paciente mesmo no começo da doença.

O uso de técnicas de *deep learning* para classificar sinais de voz de pacientes com Parkinson e diferenciar de indivíduos saudáveis é uma abordagem promissora para o diagnóstico precoce da doença, podendo ajudar a identificar padrões nos sinais de voz que detectarão a doença de Parkinson.

Com isso o trabalho tem como justificativa contribuir para o desenvolvimento de uma técnica mais precisa e eficaz para a detecção precoce da doença de Parkinson, utilizando técnicas de *deep learning* para a classificação de sinais de voz. Podendo ter um impacto significativo na melhoria da qualidade de vida dos pacientes e na redução dos custos de tratamento associados à doença.

1.3 OBJETIVOS

O objetivo principal do trabalho é detectar indivíduos com doença de Parkinson diferenciando de indivíduos sem a patologia.

É esperado que o resultado deste trabalho possa contribuir significativamente para o desenvolvimento de novas técnicas para a identificação precoce do Parkinson e, conseqüentemente, para um tratamento mais eficaz e melhor qualidade de vida para os pacientes.

1.3.1 OBJETIVOS GERAIS

Os objetivos gerais desse trabalho são aplicar técnicas de aprendizado profundo para a classificação de sinais de voz de pacientes com Parkinson e indivíduos saudáveis, além de identificar padrões nos sinais de voz que possam auxiliar no diagnóstico precoce da doença de Parkinson e avaliar a viabilidade do uso de sinais de voz para o diagnóstico do Parkinson em um ambiente clínico.

1.3.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos serão focados no intuito de atingir os objetivos geral e são as seguintes:

- Realizar uma revisão da literatura sobre a doença de Parkinson;
- Selecionar dados de datasets de sinais de voz de pacientes com Parkinson e indivíduos sem Parkinson;
- Desenvolver um modelo de *deep learning* para a classificação dos sinais de voz;

- Treinar e validar o modelo;
- Analisar os resultados e identificar padrões nos sinais de voz que possam auxiliar no diagnóstico precoce;
- Avaliar a viabilidade do uso de sinais de voz para o diagnóstico em um ambiente clínico;

1.4 METODOLOGIA DE PESQUISA

O trabalho envolve a análise de sinais de voz de pacientes com Parkinson e indivíduos saudáveis. Serão utilizadas redes neurais convolucionais para a classificação dos sinais de voz, com o objetivo de identificar padrões e características que possam ser usados como detectores da doença de Parkinson.

Esses modelos permitirão a análise de características temporais dos sinais de voz e a identificação de padrões dinâmicos que possam ser usados na classificação.

A utilização de dados será realizada por meio de datasets de pacientes com Parkinson e indivíduos saudáveis.

Serão comparados os resultados da classificação de sinais de voz obtidos com a abordagem proposta com os resultados obtidos com outras abordagens convencionais de classificação de sinais de voz.

A avaliação da eficácia da abordagem será baseada na acurácia da classificação, sensibilidade, especificidade e outras métricas de desempenho relevantes. Os resultados serão analisados e discutidos em termos de sua importância e contribuição para o diagnóstico precoce da doença de Parkinson.

2 REVISÃO TEÓRICA

2.1 DOENÇA DE PARKINSON

2.1.1 Introdução

A doença de Parkinson, uma doença neurológica de considerável impacto no sistema nervoso humano, provoca problemas notáveis na coordenação motora e nos movimentos. A sua origem reside na degeneração dos neurônios cerebrais, localizados na área conhecida como substância negra, onde a dopamina é produzida e exerce influência sobre as emoções. A diminuição dessa substância neurotransmissora, a dopamina, emerge como a principal culpada pela comprometida capacidade de movimentação. (SAÚDE, 2019)

Entre os sintomas que caracterizam a condição, destacam-se os tremores involuntários em diversas partes do corpo, como braços, pernas e cabeça, bem como a rigidez muscular, muitas vezes acompanhada de desconforto, que impacta a fluidez dos movimentos. A lentidão nas ações motoras, a dificuldade em manter a postura – levando a possíveis quedas –, passos mais curtos e a desafio em modificar a direção ao caminhar são observados frequentemente.

Além dos efeitos sobre o controle motor, o Parkinson pode influenciar a esfera emocional, propiciando o desenvolvimento de ansiedade, depressão e dificuldades de memória. Distúrbios do sono e desafios cognitivos também podem emergir em decorrência desta condição. Apesar dos avanços na pesquisa, a enfermidade permanece desprovida de um tratamento definitivo e seu mecanismo ainda não é plenamente compreendido. Sugere-se que uma interação complexa entre fatores genéticos e ambientais desempenhe um papel na degeneração neuronal, contribuindo para a variabilidade na progressão da doença entre os indivíduos afetados.

2.1.2 Sintomas

A doença de Parkinson é composto por vários sintomas que afetam principalmente a coordenação motora e o sistema nervoso. Os principais sintomas dessa doença, são mais frequentemente os tremores involuntários, que geralmente se manifestam nas mãos, braços, pernas e na cabeça. Os tremores ficam mais evidentes quando o paciente está em repouso e muitas vezes tendem a diminuir ou desaparecer durante o movimento. (SAÚDE, 2019)

Outro sintoma perceptível é a rigidez muscular, que causa resistência ao tentar movimentar as articulações, prejudicando a facilidade dos movimentos e fazendo com que seja desconfortável se movimentar. A bradicinesia, que é a lentidão nos movimentos, é também uma característica da doença. Tarefas do dia a dia que antes eram facilmente realizadas de maneira fluida podem se tornar mais lentas e de difícil execução.

A dificuldade em manter a postura é outro sintoma comum, além das dificuldades em manter o equilíbrio. Isso leva ao aumento do risco de quedas, que são observadas em pacientes com Parkinson. Além da doença poder causar alterações na maneira de caminhar, resultando em passos curtos e dificuldade para iniciar ou modificar a direção da caminhada.

Além da parte física, a doença de Parkinson também pode ter impacto na esfera emocional e cognitiva. Vários pacientes relatam alterações de humor, inclusive ansiedade e depressão, distúrbios do sono, como insônia e sonolência excessiva, dificuldade em focar a atenção, problemas de memória e dificuldades cognitivas também podem ser observados, afetando muito a qualidade de vida do paciente.

2.1.3 Detecção

Atualmente, não se dispõe de um teste definitivo para a detecção da doença de Parkinson. Como resultado, os profissionais médicos empregam uma abordagem multifacetada, utilizando uma variedade de técnicas para identificar a enfermidade o mais cedo possível. Esse processo inclui avaliações clínicas minuciosas, exames neurológicos abrangentes e procedimentos de diagnóstico complementares para descartar condições que possam compartilhar sintomas semelhantes.

No estágio inicial da avaliação, os históricos médicos e exames clínicos são cuidadosamente analisados para coletar informações cruciais sobre a manifestação e a evolução dos sintomas da doença. Os sintomas relacionados ao sistema motor são particularmente examinados, enquanto esforços são direcionados para excluir outras doenças com características clínicas comparáveis. Além disso, a administração de medicamentos pode ser explorada como uma ferramenta de diagnóstico, visando observar se há uma melhoria nos sintomas de Parkinson em resposta a medicamentos conhecidos por aliviar esses sintomas. Para ampliar a análise diagnóstica, recorre-se também a exames de imagens e testes neuropsicológicos.

Dentre os procedimentos de imagem frequentemente utilizados, encontra-se a ultrassonografia transcraniana, um exame de ultrassom realizado no crânio para identificar alterações de coloração na substância negra. Além disso, a cintilografia cerebral é empregada para avaliar os níveis de dopamina no estriado, região onde a dopamina é liberada após sua produção na substância negra (FONOFF, 2020), oferecendo assim um panorama mais abrangente da condição neurológica.

2.1.4 Coleta de Dados

A coleta de dados sobre a doença de Parkinson envolve uma gama de métodos e abordagens para obter informações detalhadas sobre os sintomas, a progressão da doença, entre outros. (SAÚDE, 2022)

Alguns dos principais meios de coleta de dados incluem o histórico médico e entrevistas, o processo começa pela obtenção de históricos médicos detalhados do paciente, incluindo informações sobre os sintomas, sobre a duração e evolução dos mesmos ao longo do tempo. Entrevistas com o paciente e seus familiares também podem fornecer informações sobre os sintomas iniciais e quaisquer mudanças.

Outra forma de realizar a coleta é a avaliação clínica por médicos especializados em neurologia, que conduzem avaliações clínicas minuciosas para observar os sintomas motores e não-motores associados à essa doença. Isso pode envolver alguns testes para avaliar tremores, rigidez muscular, bradicinesia e instabilidade postural, ou seja, os sintomas mais comuns.

Os exames neurológicos são exames físicos realizados por neurologistas ou especialistas em doença de Parkinson que podem incluir testes que avaliam a coordenação motora, a função muscular e a resposta a estímulos sensoriais que ajudam a complementar os dados.

Existem também os exames complementares que são vários exames que são empregados para apoiar o diagnóstico e avaliação da doença. Esses exames podem incluir ressonância magnética e tomografia computadorizada do cérebro para descartar outras doenças com sintomas parecidos, como tumores ou lesões cerebrais.

Testes de imagem, como ultrassonografia transcraniana e cintilografia cerebral, são realizados para analisar a atividade e a integridade das regiões cerebrais afetadas pela doença. Dispositivos de monitoramento que são possíveis de trajar, como acelerômetros, podem ser usados para registrar os movimentos ao longo do dia, fornecendo dados mais detalhados dos sintomas motores e suas flutuações.

Avaliações Psicossociais podem ajudar a avaliar as funções cognitivas, memória, habilidades de atenção e aspectos emocionais, como ansiedade e depressão, que também são sintomas da doença.

Alguns pacientes são instruídos a manter diários onde registram a intensidade dos sintomas, flutuações, momentos de maior dificuldade e fatores desencadeantes para auxiliar no monitoramento dos sintomas.

Dados de pacientes coletados de várias fontes, incluindo consultórios médicos, hospitais e instituições de pesquisa, podem ser agregados para análises epidemiológicas e de pesquisa.

E estudos de acompanhamento de longo prazo, nos quais pacientes são monitorados ao longo de anos, permitem observar a progressão da doença e suas variações.

A combinação de todos esses meios de coleta de dados ajudam os médicos e pesquisadores a obterem uma compreensão mais completa da doença de Parkinson, desde seus sintomas iniciais até o acompanhamento contínuo da progressão da doença de forma completa.

2.1.5 Necessidade de métodos de IA para detecção de Parkinson

A utilização de métodos de Inteligência Artificial para detectar a doença de Parkinson apresenta uma série de vantagens para a sua aplicação. A capacidade da IA em analisar grandes conjuntos de dados de pacientes habilita a possibilidade de identificar padrões de difícil percepção que poderiam escapar à observação humana de médicos. Com isso, a utilização desses métodos podem resultar em diagnósticos mais precisos e em estágios mais iniciais da doença.

Um dos pontos que mais se destacam é a capacidade de detectar precocemente a doença. Ao analisar de forma minuciosa os padrões de movimento e outros sintomas, a inteligência artificial pode detectar pequenas mudanças que podem sinalizar o início da doença em fases muito precoces, permitindo tratamentos antecipados e muitas vezes mais eficazes.

A grande complexidade dos dados históricos, clínicos, exames de imagem, médicos, e registros dos movimento dos pacientes exigem uma técnica que ultrapasse a capacidade humana. A inteligência artificial, já que conta com um processamento rápido e é capaz de realizar uma análise mais aprofundada, é capaz de examinar todos esses tipos de dados e fornecer ideias mais abrangentes sobre o diagnóstico do paciente. (ZAPAROLLI, 2022)

Além disso, a inteligência artificial pode ser uma boa ferramenta para personalizar os tratamentos de Parkinson. Tendo base nos dados individuais de cada paciente tratado, a IA pode ajudar na adaptação do tratamento, levando em consideração a gravidade dos sintomas e a resposta a medicamentos.

O constante monitoramento é outra área em que a IA pode ser impactante. Através de dispositivos equipáveis e sensores que coletam os dados em tempo real sobre os movimentos e sintomas dos pacientes, a IA pode realizar uma análise em tempo real, permitindo um acompanhamento mais preciso da evolução da doença ao longo do tempo.

Outras vantagens são a capacidade da inteligência artificial identificar subtipos da doença com base nos dados coletados e em características clínicas e genéticas, melhorar a pesquisa médica graças a descoberta de correlações e associações nos dados, e fornecer análises iniciais ou diagnósticos em áreas com limitada disponibilidade de especialistas.

A aplicação da IA na detecção da doença de Parkinson proporciona um grande conjunto de benefícios, incluindo diagnósticos mais precisos e precoces, personalização de tratamentos, análise de dados complexos e ideias que podem melhorar a compreensão da doença e a qualidade de vida dos pacientes.

2.2 MACHINE LEARNING

2.2.1 Redes Neurais

Redes neurais são métodos de Inteligência Artificial, feita com modelos computacionais inspirados na estrutura e funcionamento do cérebro humano, mais especificamente no modo como os neurônios se comunicam e processam informações. Essas redes são usadas em Inteligência Artificial e Aprendizado de Máquina para realizar tarefas de grande complexidade, como reconhecimento de padrões, classificação, previsão e tomada de decisões, imitando a maneira como o cérebro humano processa informações.

Uma rede neural artificial é composta por várias camadas de unidades interconectadas, conhecidas como "neurônios artificiais", que simulam as interações biológicas realizadas pelos neurônios do cérebro humano. Cada neurônio artificial recebe entradas ponderadas, realiza uma soma ponderada dessas entradas e, em seguida, aplica uma função de ativação para determinar sua saída.(SERVICES, 2023b)

Existem três tipos principais de camadas que compõem as redes neurais artificiais, são elas:

- Camada de Entrada: Nessa camada que são recebidos os dados iniciais ou entradas para a rede neural. Cada neurônio artificial na camada de entrada corresponde a um atributo ou característica dos dados.
- Camadas Intermediárias ou Ocultas: Essas camadas, como o nome indica, ficam entre a camada de entrada e a camada de saída. Cada neurônio artificial nessas camadas realiza cálculos complexos que combinam e transformam as informações das camadas anteriores.
- Camada de Saída: Esta é a camada final que produz a saída da rede após o processamento das camadas intermediárias. A saída pode ser um valor contínuo ou uma classe em um problema de classificação.

A aprendizagem em redes neurais artificiais é alcançada por meio de um processo de treinamento. Durante o treinamento, a rede é cerregada com um conjunto de dados de treinamento com entradas conhecidas e saídas desejadas. A rede neural ajusta os pesos e os parâmetros internos das conexões entre os neurônios para minimizar a diferença entre as saídas previstas e as saídas reais dos dados de treinamento. Isso é frequentemente realizado usando algoritmos de otimização, como o Gradiente Descendente.

2.2.1.1 Redes Neurais Multicamada

Redes neurais multicamadas são uma modelos avançados de aprendizado de máquina inspirados na estrutura do sistema nervoso humano, mais especificamente nos neurônios do cérebro. Essas redes consistem em várias camadas interconectadas de neurônios artificiais, permitindo um processamento mais complexo e uma capacidade maior de aprendizado do sistema. Cada camada cumpre um papel específico no processamento e transformação dos dados de entrada, tornando essas redes adequadas para tarefas mais complexas, como reconhecimento de padrões e tomada de decisões.(IYODA, 2000)

A estrutura das redes neurais multicamadas é composta por três tipos de camadas: a camada de entrada, as camadas intermediárias (ou ocultas) e a camada de saída. A camada de entrada recebe os dados iniciais que serão processados. As camadas ocultas, geralmente compostas por várias camadas, realizam cálculos complexos para transformar e combinar as informações. Cada neurônio artificial nas camadas intermediárias processa as entradas com pesos pré determinados, aplicando funções de ativação para produzir saídas intermediárias. E por fim, a camada de saída fornece a resposta final da rede neural após a conclusão do processamento.

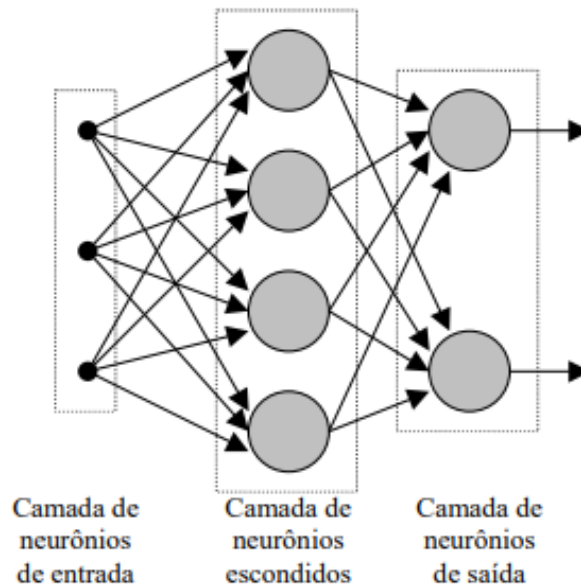


Figura 1 – Redes Neurais em Camadas (IYODA, 2000)

O treinamento dessas redes neurais multicamadas é um processo iterativo em que a rede é exposta a um conjunto de dados de treinamento com entradas conhecidas e saídas desejadas. Durante o treinamento, os pesos das interações entre os neurônios são ajustados gradualmente para diminuir a diferença entre as saídas previstas e as saídas reais do conjunto de treinamento. Isso é geralmente alcançado por algoritmos de otimização, como o Gradiente Descendente, que buscam encontrar os valores de peso que reduzem o erro da rede.

Vamos considerar uma rede neural com L camadas, onde a camada l tem n_l neurônios. Suponha que $a_j^{(l)}$ seja a ativação (saída) do j -ésimo neurônio na l -ésima camada. As ativações são calculadas da seguinte forma:

$$z_j^{(l)} = \sum_{i=1}^{N_{l-1}} w_{ji}^{(l)} a_i^{(l-1)} + b_j^{(l)} \quad (1)$$

$$a_j^{(l)} = g(z_j^{(l)}) \quad (2)$$

Onde:

- $z_j^{(l)}$ é a entrada ponderada do j -ésimo neurônio na l -ésima camada.

- $w_{ji}^{(l)}$ é o peso associado à conexão entre o i-ésimo neurônio na l-1-ésima camada e o j-ésimo neurônio na l-ésima camada.
- $b_j^{(l)}$ é o termo de polarização (bias) associado ao j-ésimo neurônio na l-ésima camada.
- g é uma função de ativação aplicada ao resultado ponderado $z_j^{(l)}$

A última camada de (L) é a camada de saída, e dependendo da rede neural, pode haver diferentes funções de ativação e critérios de perda associados a ela.

As redes neurais multicamadas são adequadas para resolver problemas complexos que envolvem dados de alta dimensionalidade e padrões não lineares. Já que possui a capacidade de aprender representações hierárquicas de dados, essas redes têm sido fundamentais em aplicações como reconhecimento de imagens, processamento de linguagem natural, análise de dados biomédicos e muito mais. Elas representam uma evolução significativa no campo da Inteligência Artificial, capacitando sistemas para realizar tarefas cada vez mais sofisticadas e precisas.

2.2.1.2 Redes Neurais Convolucionais

Redes Neurais Convolucionais (CNNs) são uma classe de arquiteturas de redes neurais profundas projetadas especificamente para tarefas de processamento de imagens e análise de dados espaciais. Inspiradas pela organização do sistema visual do ser humano, as CNNs demonstraram um desempenho excepcional em tarefas de visão computacional, como detecção de objetos, segmentação de imagens e reconhecimento de padrões visuais.(IBM, 2023a)

O principal componente que distingue as CNNs é a camada de convolução, que aplica operações de convolução às entradas da rede. A convolução envolve o uso de filtros (*kernels*) que percorrem a imagem em passos determinados, calculando a combinação ponderada dos pixels vizinhos. Isso permite à rede capturar características locais e padrões relevantes, como bordas, texturas e formas, de maneira eficaz.

A operação de convolução pode ser descrita da seguinte forma:

$$z(i, j) = (X * W)(i, j) = \sum_m \sum_n X(m, n) * W(i - m, j - n) \quad (3)$$

Onde:

- $Z(i,j)$ é o elemento na posição (i,j) da saída.
- $X(m,n)$ é o elemento na posição (m,n) da entrada.
- $W(i,j)$ é o elemento na posição (i,j) do filtro.

Além das camadas de convolução, as CNNs também podem incluir camadas de agrupamento (ou *pooling*), que reduzem a dimensão espacial das representações, mantendo as características mais relevantes. Isso resulta em uma redução na quantidade de parâmetros e no consumo de recursos computacionais, ao mesmo tempo que preserva a informação essencial.

A operação de *pooling* pode ser descrita da seguinte forma em um janela de tamanho $f \times f$:

$$Y(i, j) = \max_{m,n} X(f * i + m, f * j + n) \quad (4)$$

Onde:

- $Y(i,j)$ é o elemento na posição (i,j) da saída.
- $X(f*i+m,f*j+n)$ é o elemento na posição $(f*i+m,f*j+n)$ da entrada.
- f é o tamanho da janela de *pooling*.

Uma característica importante das CNNs é a capacidade de aprendizado hierárquico de características. À medida que as camadas de convolução e agrupamento são empilhadas, a rede é capaz de aprender representações cada vez mais complexas e abstratas das imagens de entrada. Isso é muito útil para tarefas de visão computacional, onde as características visuais são organizadas em diferentes níveis de abstração.

O treinamento das CNNs é realizado por meio da apresentação de um grande conjunto de dados de treinamento, em que os pesos das conexões são ajustados para minimizar o erro entre as saídas previstas e as saídas reais. A popularidade das CNNs cresceu significativamente devido à sua eficácia em muitas tarefas de análise de imagens e à capacidade de capturar automaticamente características relevantes, eliminando a necessidade de extração manual de características.

As CNNs ,hoje em dia , são amplamente empregadas em áreas como reconhecimento facial, detecção de objetos, diagnóstico médico por imagem e veículos autônomos, desempenhando um papel crucial na expansão das capacidades de sistemas de visão computacional e em aplicações práticas que envolvem análise de dados visuais complexos.

2.2.1.3 Redes Neurais Recorrentes

As RNNs são uma classe de modelos de aprendizado de máquina que são eficazes para lidar com dados sequenciais, como séries temporais, texto e áudio. Elas são projetadas para capturar relações temporais em dados, o que as torna poderosas em uma ampla gama de aplicações.(JONES, 2017)

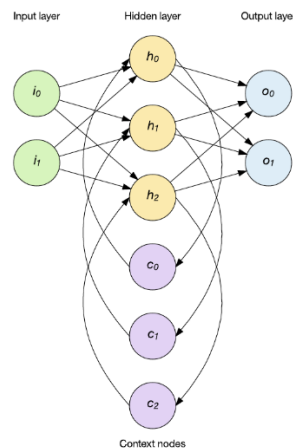


Figura 2 – Redes Neurais Recorrentes
(JONES, 2017)

A principal característica das RNNs é que elas têm conexões retroativas em seu próprio grafo, permitindo que informações anteriores influenciem nas previsões atuais. A inovação é a introdução de um ciclo temporal, que permite que informações sejam mantidas e atualizadas ao longo do tempo. Isso é essencial para lidar com sequências de comprimento variável.

As RNNs são capazes de processar sequências de entrada de comprimento aleatório. Elas percorrem a sequência em passos de tempo sucessivos, atualizando seu estado interno com base nas entradas anteriores.

As RNNs mantêm um estado oculto que é atualizado a cada passo de tempo. Este estado oculto serve como uma representação interna que contém informações sobre o histórico da sequência.

O estado oculto h_t em um determinado passo de tempo t é calculado a partir da entrada atual (x_t) e do estado oculto anterior (h_{t-1}) usando uma função de ativação (f):

$$h_t = f(W_{hx} + W_{hh} * h_{t-1} + b_h) \quad (5)$$

Onde:

- W_{hx} é a matriz de pesos que mapeia a entrada para o estado oculto.
- W_{hh} é a matriz de pesos que mapeia o estado oculto anterior para o estado atual.
- b_h é o vetor de *bias*.
- f é uma função de ativação, geralmente uma função não linear como a tangente hiperbólica ou a função sigmóide.

A retropropagação através do tempo é o algoritmo de treinamento usado para RNNs. Ele é uma extensão do algoritmo de retropropagação tradicional que leva em conta a natureza sequencial dos dados.

As RNNs podem enfrentar problemas de explosão do gradiente durante o treinamento. Isso ocorre quando os gradientes se tornam muito pequenos ou muito grandes à medida que são propagados ao longo do tempo. Estratégias como *gradient clipping* e arquiteturas mais avançadas, como LSTMs e GRUs, foram desenvolvidas para lidar com esse problema.

As LSTMs são uma variante avançada de RNNs que foram projetadas para capturar dependências temporais mais longas. Elas introduzem mecanismos de "portões" que controlam o fluxo de informações no estado oculto, permitindo que informações relevantes sejam mantidas por longos períodos de tempo.

As GRUs são outra variante de RNN que simplificam a arquitetura das LSTMs, mantendo a capacidade de capturar dependências temporais de longo prazo.

2.2.2 Deep Learning

O aprendizado profundo é uma subárea do aprendizado de máquina que se concentra em treinar modelos de rede neural para aprender representações hierárquicas de dados. Ele tem sido extremamente bem-sucedido em uma ampla variedade de tarefas, desde reconhecimento de imagem até processamento de linguagem natural e muito mais. (ORACLE, 2023)

O aprendizado profundo pode ser utilizado tanto como aprendizados supervisionados quanto não supervisionados. No aprendizado supervisionado, os modelos são treinados em pares de entrada e saída rotulados. No aprendizado não supervisionado, os modelos aprendem padrões nos dados sem rótulos, podendo ser utilizados em tarefas como agrupamento e redução de dimensionalidade.

O treinamento de redes neurais profundas geralmente envolve a minimização de uma função de custo usando algoritmos de otimização como o gradiente descendente e suas variantes. O *backpropagation* é a técnica chave que calcula os gradientes dos pesos em relação à função de custo, permitindo a atualização iterativa dos pesos durante o treinamento.

O aprendizado profundo tem sido utilizado em uma variedade de aplicações como reconhecimento de imagem e visão computacional para a classificação de objetos em imagens, detecção de objetos, segmentação de imagens, entre outros, também é utilizado em processamento de linguagem natural, reconhecimento de fala, jogos, robótica e biomedicina com os diagnósticos médicos, análises de imagens médicas, entre outros.

O aprendizado profundo tem revolucionado muitos campos e continua a ser uma área de pesquisa e desenvolvimento muito ativa. Sua capacidade de aprender representações complexas a partir de dados brutos tem contribuído muito para avanços significativos em IA e aprendizado de máquina.

2.2.3 Attention

O mecanismo de atenção é uma técnica importante em aprendizado de máquina que tenta imitar a atenção cognitiva humana, é muito útil especialmente em tarefas de processamento de linguagem natural e de visão computacional. Esse mecanismo permite que os modelos deem mais importância a certas partes da entrada durante o processo de inferência, tornando-os mais flexíveis e capazes de lidar com sequências de tamanhos variáveis.

Em processamento de linguagem natural utilizam-se modelos sequência-para-sequência (seq2seq) que são uma classe especial de arquitetura das RNR. Normalmente são usados para resolver problemas como tradução, *chat-bots* e resumo de textos. São formados por uma arquitetura *encoder-decoder*, onde o encoder processa a entrada e comprime a informação em um vetor de tamanho fixo e o decoder então faz a tradução, porém há um grande problema com esse método que é as RNRs usarem tamanhos fixos para entrada e saída, dando problema quando uma tradução tem mais ou menos palavras, para resolver isso o mecanismo de atenção foca em diferentes partes da entrada para gerar a saída. Durante cada passo de tempo, o decodificador calcula a atenção sobre os estados ocultos do codificador para determinar quais partes da sequência de entrada são mais relevantes para gerar a próxima palavra.(DANTAS, 2019)

O calculo de atenção é feito com a seguinte função *softmax*:

$$a_t(i) = \frac{\exp(e_t(i))}{\sum_{j=1}^n \exp(e_t(j))} \quad (6)$$

Onde:

- $e_t(i)$ é uma pontuação de atenção calculada usando a relação entre o estado oculto atual do decodificador e o estado oculto i do codificador.

O vetor de contexto é calculado pela soma ponderada dos estados ocultos em relação aos seus pesos. Esse vetor é calculado para cada palavra na saída:

$$c_t = \sum_{i=1}^n a_t(i) * h_i \quad (7)$$

Com o vetor de contexto juntamente com o estado oculto resulta em uma saída correta.

2.2.4 Transformers

Os *Transformers* são uma arquitetura de modelo de aprendizado de máquina introduzida por Ashish Vaswani e seus colegas em um artigo chamado "*Attention is All You Need*" em 2017. Eles revolucionaram muitas áreas de processamento de linguagem natural e visão computacional e têm sido amplamente adotados em uma variedade de aplicações.

O ponto principal dos *Transformers* é o mecanismo de atenção, que permite ao modelo ponderar a importância de diferentes partes da entrada durante a geração de uma saída. Os *Transformers* não se limitam a processar a entrada sequencialmente, mas são capazes de considerar todas as interações entre os elementos da sequência simultaneamente. Isso resulta em uma alta eficiência computacional e na capacidade de lidar com sequências de comprimento variável com facilidade.(MAXIME, 2019)

Uma das inovações dos *transformers* é a autoatenção, onde cada elemento na sequência é associado a um peso que representa sua relevância para todos os outros elementos na mesma sequência. Essa abordagem, combinada com o uso de múltiplos focos de atenção, permite ao modelo aprender diferentes tipos de atenção e capturar relações complexas entre os elementos.

A arquitetura de um *transformer* é composta por várias camadas empilhadas. Cada camada contém uma subcamada de atenção, responsável por calcular os pesos de atenção, e uma subcamada de *feedforward*, que adiciona um componente não linear ao modelo. Camadas de normalização residual e conexões residuais são usadas para melhorar a estabilidade do treinamento, permitindo um fluxo mais suave de gradientes durante a retropropagação.

Os *transformers* demonstraram serem poderosos em diversas aplicações, desde tradução automática e sumarização de texto até resposta a perguntas e tarefas de visão computacional. Modelos pré-treinados baseados em *transformers*, como o BERT da *google* e o chat GPT da *open AI*, atingiram resultados de referência na área em várias tarefas de PLN.

2.2.5 Support Vector Machines

As *Support Vector Machines* são um algoritmo de aprendizado de máquina utilizado tanto para tarefas de classificação quanto de regressão. Elas são especialmente eficazes em problemas de classificação em que a separação entre as classes é complexa ou não linear.

A ideia central por trás das SVMs é encontrar o hiperplano de decisão que melhor separa as diferentes classes de dados. O hiperplano é uma superfície que divide o espaço de características em duas regiões distintas, uma para cada classe. O objetivo é encontrar o hiperplano que maximiza a margem, ou

seja, a distância entre os pontos mais próximos de cada classe e o próprio hiperplano.(DEVELOPERS, 2023b)

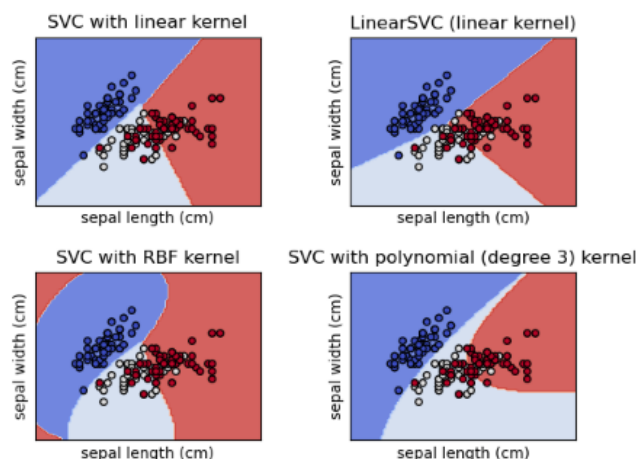


Figura 3 – Exemplo de *Support Vector Machine* (DEVELOPERS, 2023b)

A principal vantagem das SVMs é a capacidade de lidar com conjuntos de dados complexos, mesmo quando as relações entre as características não são lineares. Isso é possível através do uso de truques de *kernel*, que transformam o espaço de características para um espaço de dimensão superior onde a separação entre as classes pode ser feita de forma mais eficaz. Além disso, as SVMs são bastante eficazes em lidar com conjuntos de dados de alta dimensionalidade, sendo menos propensas a *overfitting* em comparação com outros algoritmos.

Existem diferentes tipos de SVMs, incluindo SVMs de classificação binária, SVMs de classificação multi-classe e SVMs de regressão. Além disso, existem várias variantes de SVMs, como as SVMs com margem suave que permitem algum erro de classificação no treinamento para lidar com conjuntos de dados que não são linearmente separáveis.

2.2.6 Decision Tree

Uma Árvore de Decisão é um algoritmo de aprendizado de máquina utilizado em tarefas de classificação e regressão. Ela representa um modelo preditivo na forma de uma estrutura de árvore, onde cada nó interno representa uma decisão baseada em uma característica do conjunto de dados. Os nós folha representam as saídas ou classes finais.(DEVELOPERS, 2023a)

O processo de construção de uma Árvore de Decisão envolve a seleção das características que melhor dividem os dados em subgrupos homogêneos. Isso é feito de maneira recursiva, dividindo o conjunto de dados em subconjuntos cada vez mais puros até que critérios de parada sejam atendidos.

Uma das vantagens das Árvores de Decisão é a sua capacidade de lidar com conjuntos de dados heterogêneos, incluindo dados categóricos e numéricos. No entanto, as Árvores de Decisão também têm algumas limitações. Elas podem ser propensas a *overfitting*, especialmente quando não são adequadamente podadas. Além disso, em alguns casos, podem ser sensíveis a pequenas variações nos dados de treinamento.

Para mitigar essas preocupações, técnicas como a poda da árvore, o uso de Random Forest e a implementação de critérios de parada bem escolhidos são frequentemente aplicadas.

2.2.7 Random Forest

O *Random Forest* é um algoritmo de aprendizado de máquina que se destaca em problemas de classificação e regressão.(IBM, 2023b)

Ele é construído a partir de um conjunto de árvores de decisão. Cada árvore é treinada em uma amostra aleatória do conjunto de dados e utiliza um subconjunto aleatório das características disponíveis. Isso ajuda a reduzir a correlação entre as árvores e a aumentar a diversidade do modelo. A classificação final é obtida através da combinação das saídas de todas as árvores individuais. Para problemas de classificação, é feita uma votação entre as árvores para determinar a classe mais provável, enquanto em problemas de regressão, é calculada a média das previsões das árvores.

Devido à diversidade e aleatoriedade na construção das árvores, o *Random Forest* é menos suscetível a *overfitting* do que uma única árvore de decisão. Ele tende a generalizar bem para dados não vistos. Além disso fornece uma medida de importância para cada característica no processo de classificação. Isso pode ser útil para entender quais variáveis têm maior influência no modelo.

O *Random Forest* é amplamente utilizado em diversas áreas, incluindo classificação de imagem, detecção de fraudes, bioinformática, processamento de linguagem natural, entre outras. Embora o *Random Forest* tenha um bom desempenho com os parâmetros padrão, é possível ajustar hiperparâmetros para otimizar ainda mais o desempenho do modelo.

2.2.8 Regressão Logística

A regressão logística é uma técnica estatística essencial no campo da análise de dados, especialmente quando lidamos com problemas de classificação binária. Ao contrário da regressão linear, que prevê valores contínuos, a regressão logística é usada para prever a probabilidade de um evento ocorrer, geralmente expressa em termos de "sim" ou "não", "1" ou "0".(SERVICES, 2023a)

Essa técnica é particularmente valiosa em situações em que a variável de interesse é categórica e binária, como por exemplo, determinar se um paciente tem uma determinada condição médica ou não, com base em uma série de características ou fatores de risco.

A regressão logística utiliza uma função logística para modelar a relação entre as variáveis independentes e a probabilidade de o evento ocorrer. Essa função transforma a saída da regressão linear em uma escala de 0 a 1, representando a probabilidade.

Além disso, a regressão logística também nos fornece métricas de avaliação do modelo, como a acurácia, precisão, *recall*. Essas métricas são cruciais para avaliar o desempenho do modelo e determinar sua utilidade na prática.

2.2.9 XGBoost

O *XGBoost*, abreviação de *eXtreme Gradient Boosting* se destaca em problemas de classificação e regressão. Ele é uma implementação otimizada do algoritmo de *Gradient Boosting*, que é uma técnica de *ensemble learning*, que constrói um modelo preditivo a partir de um conjunto de modelos mais simples para melhorar a precisão da previsão.(DEVELOPERS, 2022)

O modelo incorpora técnicas de regularização para evitar *overfitting*. Ele utiliza termos de penalização na função de custo, o que ajuda a controlar a complexidade do modelo e a reduzir a chance de sobreajuste aos dados de treinamento. Ele utiliza uma função de custo que combina termos de erro de previsão e termos de regularização. Isso permite a otimização do modelo de acordo com os objetivos do problema, como minimizar o erro de classificação ou minimizar o erro quadrático médio.

Ele tem a capacidade de aprender automaticamente como lidar com valores ausentes nas características durante o treinamento. Uma das grandes vantagens do *XGBoost* é sua eficiência computacional. Ele é projetado para ser altamente paralelizável, o que significa que pode treinar modelos rapidamente em grandes conjuntos de dados.

2.2.10 Naive Bayes

O *Naive Bayes* é um algoritmo de aprendizado de máquina utilizado em tarefas de classificação. Ele se baseia no teorema de Bayes, que descreve a probabilidade de um evento com base em informações prévias relacionadas a esse evento. O "*naive*" (ingênuo) no nome do algoritmo refere-se à suposição de independência entre as características do conjunto de dados, o que simplifica o cálculo das probabilidades condicionais.(GUSMÃO, 2023)

Ele é especialmente eficaz em lidar com conjuntos de dados de alta dimensão e é frequentemente utilizado em tarefas de processamento de linguagem natural, como classificação de texto, filtragem de spam e análise de sentimentos. Ele também é aplicado em problemas de aprendizado supervisionado em geral, como classificação de e-mails, diagnóstico médico, detecção de fraudes e muito mais.

A força do *Naive Bayes* reside na sua simplicidade e eficiência computacional. Ele requer um tempo de treinamento relativamente curto e pode lidar bem com grandes volumes de dados. No entanto, a suposição de independência entre as características nem sempre é realista em cenários do mundo real, o que pode afetar a precisão do modelo.

Apesar dessa simplificação, o *Naive Bayes* muitas vezes produz resultados surpreendentemente bons em muitos tipos de conjuntos de dados. É uma ferramenta valiosa em aprendizado de máquina, especialmente quando se lida com conjuntos de dados textuais e situações em que a eficiência computacional é uma consideração importante.

2.2.11 K-Neighbors

O algoritmo *k-Nearest Neighbors* é uma técnica em aprendizado de máquina, usada para classificação e regressão. Sua abordagem é intuitiva e baseada na ideia de que objetos semelhantes tendem a estar próximos uns dos outros no espaço de características.(GEEKSFORGEEKS, 2023)

A essência do k-NN reside na sua simplicidade. Para classificar um novo ponto de dados, o algoritmo examina os k pontos de dados mais próximos no conjunto de treinamento. A classe mais frequente entre esses vizinhos é atribuída ao ponto de dados em questão. No caso da regressão, a média ou mediana dos valores dos k vizinhos é tomada como a previsão.

A escolha do valor de k é crucial. Um valor muito pequeno pode tornar o modelo sensível a ruídos nos dados, enquanto um valor muito grande pode resultar em uma generalização excessiva, levando a previsões menos precisas.

O k-NN é flexível e pode lidar com diferentes tipos de dados. No entanto, ele pode ser computacionalmente custoso, pois precisa calcular a distância entre o ponto de teste e todos os pontos de treinamento.

O k-NN é especialmente útil em situações onde a distribuição dos dados pode ser complexa ou não linear, e quando não há suposições claras sobre a natureza subjacente do problema.

3 REVISÃO BIBLIOGRAFICA

3.0.1 Inteligência Artificial e Medicina

Nesse primeiro artigo foi analisado como a inteligência artificial vem modificando e melhorando a medicina como um todo, proporcionando diagnósticos rápidos e auxiliando médicos no seu trabalho. (LOBO, 2017)

Atualmente, a relação entre médico e paciente, assim como a eficácia do exame clínico, estão sendo debatidas. A dependência de exames complementares para diagnósticos clínicos tem se tornado mais evidente. Paralelamente, cresce a importância da tecnologia na medicina e saúde pública. Isso se reflete na adoção de sistemas de apoio à decisão clínica, na integração de novas tecnologias, como os dispositivos vestíveis, e no armazenamento em grande escala de dados de saúde. Esse fenômeno é conhecido como big data.

A Inteligência Artificial desempenha um papel crucial nesse cenário, processando esses dados por meio de algoritmos que, por sua vez, tendem a aprimorar seu desempenho por meio do aprendizado próprio. Dessa forma, são capazes de oferecer diagnósticos cada vez mais precisos.

Exemplos marcantes incluem o supercomputador *Watson* da IBM, que compila vasto conhecimento médico e é consultado por especialistas ao redor do mundo. Além disso, o *Deep Mind*, da *Google*, registrou dados de 1,6 milhão de pacientes no Serviço Nacional de Saúde do Reino Unido, permitindo o desenvolvimento de novos sistemas de apoio à decisão clínica. Esses sistemas analisam os dados dos pacientes e emitem alertas sobre sua evolução, evitando prescrições de medicamentos contraindicados ou conflitantes, e mantendo os profissionais de saúde atualizados sobre seus pacientes.

O *Deep Mind* também se destacou ao avaliar imagens dermatológicas na pesquisa de melanoma, superando a performance de especialistas. Isso evidencia o potencial das tecnologias para aprimorar a precisão dos diagnósticos.

No entanto, é fundamental ressaltar que, embora o computador forneça informações sobre o que está ocorrendo, cabe ao médico dialogar com o paciente sobre seu problema de saúde e as possíveis soluções, explicando o porquê de cada decisão. Isso demanda uma educação médica contínua, com ênfase na compreensão da fisiopatologia dos processos orgânicos e no desenvolvimento das habilidades de escuta, exame e orientação ao paciente, permitindo o diagnóstico e tratamento adequados, assim como o acompanhamento da evolução do quadro de saúde.

3.0.2 A new approach: information gain algorithm-based k-nearest neighbors hybrid diagnostic system for Parkinson's disease

Nesse artigo descreve sobre o desenvolvimento de sistemas especializados capazes de diagnosticar automaticamente e com alta precisão estágios iniciais da doença de Parkinson com base em sinais de fala. Foi desenvolvido um modelo baseado no Algoritmo de Ganho de Informação *K-Nearest Neighbors* para essa detecção. (YÜCELBAŞ, 2021)

Essa abordagem foi aplicada a conjuntos de dados de características formados usando o método de Transformada *Wavelet* do Fator Q Sintonizável. Inicialmente, doze subconjuntos de dados de

características formando o grupo de características desse método foram analisados separadamente, e o melhor desempenho foi selecionado. Em seguida, o modelo *K-Nearest* foi aplicado a este subconjunto de dados de características. Por fim, observou-se que os resultados de desempenho fornecidos pelo sistema para este subconjunto de dados foram superiores aos obtidos com o conjunto completo de dados.

De acordo com os resultados, os valores da curva de operação do receptor e das curvas de precisão-*recall* excederam 0,95, e uma precisão de classificação de quase 98% foi obtida com as 22 características selecionadas deste subgrupo. Além disso, o coeficiente kappa foi de 0,933, indicando um acordo perfeito entre os valores reais e previstos. O desempenho do sistema *K-Nearest* também foi comparado com resultados de outros estudos na literatura que utilizaram os mesmos dados, e a abordagem proposta neste estudo superou significativamente qualquer outra abordagem relatada na literatura. Além disso, assim como nesta abordagem, um sistema especializado capaz de diagnosticar a doença de Parkinson e atingir o máximo desempenho com menos características a partir dos sinais de áudio não havia sido encontrado anteriormente.

3.0.3 Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset

O objetivo deste artigo é desenvolver um método de detecção automática com base no mais recente conjunto de dados chinês coletado. Ao contrário do inglês, não houve consenso sobre as principais características que indicam distúrbios de linguagem devido a disfunção dos órgãos vocais. Portanto, uma de nossas abordagens é classificar a fonação e articulação da fala com um modelo de seleção de características baseado em aprendizado de máquina. (WANG YAN FU, 2022a)

Com base em uma amostra relativamente grande, três algoritmos de seleção de características (LASSO, mRMR, Relief-F) foram testados para selecionar as características vocais extraídas de sinais de fala coletados em um ambiente controlado, seguido por quatro classificadores, *Naïve Bayes*, *K-Nearest Neighbor*, Regressão Logística e Descida de Gradiente Estocástica para detectar o distúrbio.

A abordagem proposta apresenta uma precisão de 75,76%, sensibilidade de 82,44%, especificidade de 73,15% e precisão de 76,57%, indicando a viabilidade e um futuro promissor para a detecção automática e discreta de DP na população chinesa.

A comparação entre os três algoritmos de seleção revela que o seletor LASSO tem o melhor desempenho, independentemente dos tipos de características vocais. A melhor precisão de detecção é obtida pelo classificador SGD, enquanto a melhor sensibilidade resultante é obtida pelo classificador LR.

Mais interessante ainda, as características de articulação são mais representativas e indicativas do que as características de fonação entre todos os algoritmos de seleção e classificação. As características de articulação mais proeminentes são F1, F2, DDF1, DDF2, BBE e MFCC.

3.0.4 A generic optimization and learning framework for Parkinson disease via speech and handwritten records

O artigo propõe que o estudo atual propõe um quadro genérico para o diagnóstico da doença de Parkinson usando imagens manuscritas e sinais de fala. Para as imagens manuscritas, foram treinadas 8 redes neurais convolucionais pré-treinadas por meio de transferência de aprendizado, ajustadas pelo "Otimizador Aquila", no conjunto de dados *NewHandPD* para diagnosticar a doença de Parkinson. (YOUSIF HOSSAM MAGDY BALAHA, 2023)

Para os sinais de fala, foram extraídas características do conjunto de dados MDVR-KCL numericamente usando dezesseis algoritmos de extração de características diferentes e alimentadas a quatro algoritmos de aprendizado de máquina diferentes ajustados pelo algoritmo *Grid Search*, além de serem processadas graficamente usando cinco técnicas diferentes e alimentadas às oito estruturas CNN pré-treinadas.

Os autores propõem uma nova técnica para extrair características do conjunto de dados de voz com base na segmentação de durações variáveis de segmentos de sinal de fala, ou seja, o uso de diferentes durações na fase de segmentação. Usando a técnica proposta, foram gerados cinco conjuntos de dados com duzentas e oitenta e uma características numéricas.

Resultados de diferentes experimentos foram coletados e registrados. Para o conjunto de dados *NewHandPD*, a métrica melhor relatada é de 99,75% usando a estrutura VGG19. Para o conjunto de dados MDVR-KCL, as métricas melhores relatadas são de 99,94% usando os algoritmos de aprendizado de máquina KNN e SVM e as características numéricas combinadas; e 100% usando as características gráficas do mel-espectrograma combinado e a estrutura VGG19. Estes resultados superam pesquisas de ponta em estado-da-arte.

3.0.5 A Speech-Based Hybrid Decision Support System for Early Detection of Parkinson's Disease

Neste artigo, foi proposto um sistema híbrido de detecção de Parkinson. Dois conjuntos de dados de fala foram utilizados no desenvolvimento deste sistema: o primeiro é um conjunto de dados italiano de voz e fala na doença de Parkinson, e o segundo é o conjunto de dados de Gravações de Voz em Dispositivos Móveis do *King's College London*. (LAMBA TARUN GULATI, 2023)

Dezessete características acústicas foram geradas a partir das amostras de voz disponíveis nos conjuntos de dados usando a biblioteca Parselmouth. Além disso, com base na importância das características, as oito características mais significativas foram utilizadas no desenvolvimento do modelo.

Essas características foram selecionadas usando o método de algoritmo genético. Quatro classificadores, *K-Nearest Neighbors*, *XGBoost*, *Random Forest* e Regressão Logística, foram utilizados durante a etapa de classificação. Para a análise do sistema desenvolvido, foram utilizados parâmetros como acurácia, sensibilidade, f-medida, especificidade e precisão.

A combinação da abordagem de seleção de características baseada em algoritmo genético e o classificador de regressão logística obteve uma precisão de 100% no conjunto de dados italiano de voz e fala na doença de Parkinson. A mesma combinação de extração de características e classificador no

conjunto de dados de Gravações de Voz em Dispositivos Móveis do *King's College London* atingiu um nível de precisão de 90%. Os resultados mostraram que o sistema proposto superou o sistema encontrado na literatura.

3.0.6 Early detection of Parkinson's disease using machine learning

Este artigo destaca o uso de técnicas de aprendizado de máquina em telemedicina para detectar a doença de Parkinson em seus estágios iniciais. Foram realizadas pesquisas nos dados de áudio MDVP de 30 pessoas com DP e pessoas saudáveis durante o treinamento de quatro modelos de aprendizado de máquina. (GOVINDU, 2023)

A comparação dos resultados de classificação pelos modelos de Máquina de Vetores de Suporte, *Random Forest*, *K-Nearest Neighbors* e Regressão Logística indicou que o classificador *Random Forest* é a técnica ideal de Aprendizado de Máquina para detecção da doença.

O modelo de classificação *Random Forest* tem uma precisão de detecção de 91,83% e sensibilidade de 0,95. Tem como objetivo promover o uso de aprendizado de máquina em telemedicina, proporcionando assim uma nova perspectiva de vida para pacientes que sofrem da doença de Parkinson.

4 DATASETS

Foram empregados um total de três conjuntos de dados na análise, com o objetivo de alcançar uma conclusão acerca da eficácia da aplicação de aprendizado de máquina no campo da medicina, especificamente na detecção da doença de Parkinson.

Cada um desses conjuntos de dados possui características distintas que serão minuciosamente detalhadas a seguir, proporcionando assim um entendimento mais completo e aprofundado do estudo em questão.

4.0.1 Parkinson's Disease Data Set

No contexto deste primeiro conjunto de dados, foram registradas medições de voz de 31 indivíduos, dos quais 23 foram diagnosticados com a doença de Parkinson. Cada coluna nesse conjunto de dados representa uma medida de voz específica, enquanto cada linha corresponde a uma das 195 gravações de voz realizadas com os pacientes, totalizando aproximadamente seis gravações por indivíduo. (UKANI, 2020)

As características relevantes a esse conjunto de dados, que já foram previamente extraídas, incluem:

- Nome: Nome indentificador e o número da gravação de cada indentificador.
- MDVP:Fo(Hz): Frequência vocal média.
- MDVP:Fhi(Hz): Frequência vocal máxima.
- MDVP:Flo(Hz): Frequência vocal mínima.
- MDVP:Jitter(%), MDVP:Jitter(Abs), MDVP:RAP, MDVP:PPQ, Jitter:DDP: Diversas medidas de variação na frequência.
- MDVP:Shimmer,MDVP:Shimmer(dB),Shimmer:APQ3,Shimmer:APQ5,MDVP:APQ,Shimmer:DDA: Diversas medidas de variação na amplitude.
- NHR, HNR: Medidas de proporção entre ruído e componentes tonais na voz.
- Status: O estado de saúde do paciente sendo 1 com Parkinson e 0 saudável
- RPDE, D2: Medidas de complexidade dinâmica não linear.
- DFA: Expoente de escala fractal de sinal.
- Spread1,spread2,PPE: Medidas não lineares de variação de frequência.

4.0.2 Parkinson's Disease (PD) classification

No segundo conjunto de dados, foram reunidas informações de 188 pacientes diagnosticados com Parkinson, compreendendo 107 homens e 81 mulheres, com idades variando de 33 a 87 anos. O grupo de controle, composto por indivíduos saudáveis, inclui 64 participantes, dos quais 23 são homens e 41 são mulheres, com idades entre 41 e 82 anos. Durante a coleta, o microfone foi ajustado para uma taxa de amostragem de 44,1 KHz, e foram obtidas três repetições de áudio de cada paciente, com a finalidade de analisar a pronúncia da vogal "a".(BISWAS, 2019)

Esse conjunto de dados abrange um total de 755 colunas, representando diversas características, tais como informações de frequência temporal, coeficientes cepstrais de frequência mel (MFCCs), recursos derivados de transformada *wavelet*, características da dobra vocal e recursos TWQT. As linhas, por sua vez, correspondem às três gravações de áudio para cada um dos 252 pacientes, totalizando 756 entradas no conjunto de dados.

4.0.3 Mobile Device Voice Recordings at King's College London (MDVR-KCL) from both early and advanced Parkinson's disease patients and healthy controls

Esse conjunto de dados foi feito com um celular *smartphone* Motorola Moto G4, as gravações de voz foram desenvolvidas com um aplicativo de gravação, que utiliza as mesmas funcionalidades do módulo de gravação de voz usado no aplicativo i-PROGNOSIS Smartphone, mas implantado no Android. O serviço de captura de voz funciona como um serviço de segundo plano independente do dispositivo de gravação e aciona gravações de voz através de sinais no gancho e fora do gancho do *smartphone*. As gravações são de alta qualidade com uma taxa de 44,1 kHz e uma profundidade de bits de 16 bits.

Foi pedido ao participante para relaxar um pouco e depois fazer uma ligação para o executor do teste. Cada participante leu “O Vento Norte e o Sol” e dependendo do participante, também leu “Tech. Motor. Aplicações de computador em trecho geográfico”.

Para outro arquivo de áudio iniciou-se um diálogo espontâneo com o participante, o executor do teste começa a fazer perguntas aleatórias sobre locais de interesse, trânsito local ou interesses pessoais.

No total são 37 testes com os pacientes lendo textos, sendo 21 sem a doença e 16 com a doença de Parkinson, e 36 de diálogo espontâneo sendo 21 sem e 15 com a doença.(JAEGER DHAVAL TRIVEDI, 2019)

Os arquivos vieram em formato .wav, então para o melhor treino pelo *deep learning*, foram extraídos de todos os áudios as features: *Signals*, *Fast Fourier Transform(fft)*, *Filter Bank(fbank)* e *Mel-Frequency Cepstrum Coefficients(mfccs)*.

4.1 ANALISE DOS DADOS

A fim de obter uma compreensão mais abrangente sobre a disposição dos dados nos conjuntos de dados, inicialmente foram realizadas análises abordando uma série de características. Essas análises visaram avaliar a distribuição dos dados, verificar se ela está equitativa e investigar as relações entre as características e o diagnóstico de Parkinson.

4.1.1 Primeiro Dataset

No que se refere ao primeiro conjunto de dados, é observável que a distribuição das 195 gravações se estabeleceu em uma proporção de 75% para casos de doença de Parkinson e 25% para indivíduos saudáveis. Isso implica que não se faz necessária uma normalização do conjunto de dados para uma eficaz aplicação de *Machine Learning*.

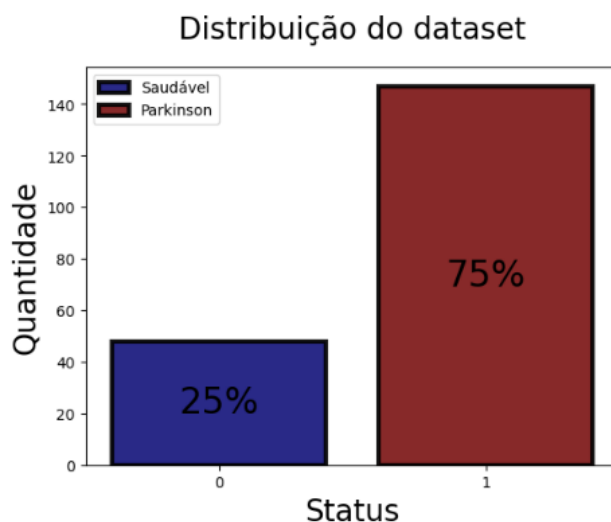


Figura 4 – Distribuição do Primeiro *Dataset*

Posteriormente, examinamos a relação de cada característica quando aplicada em indivíduos saudáveis e naqueles com a condição da doença, notando uma distinta discrepância nos picos em cada gráfico, tornando a diferença entre eles evidente.

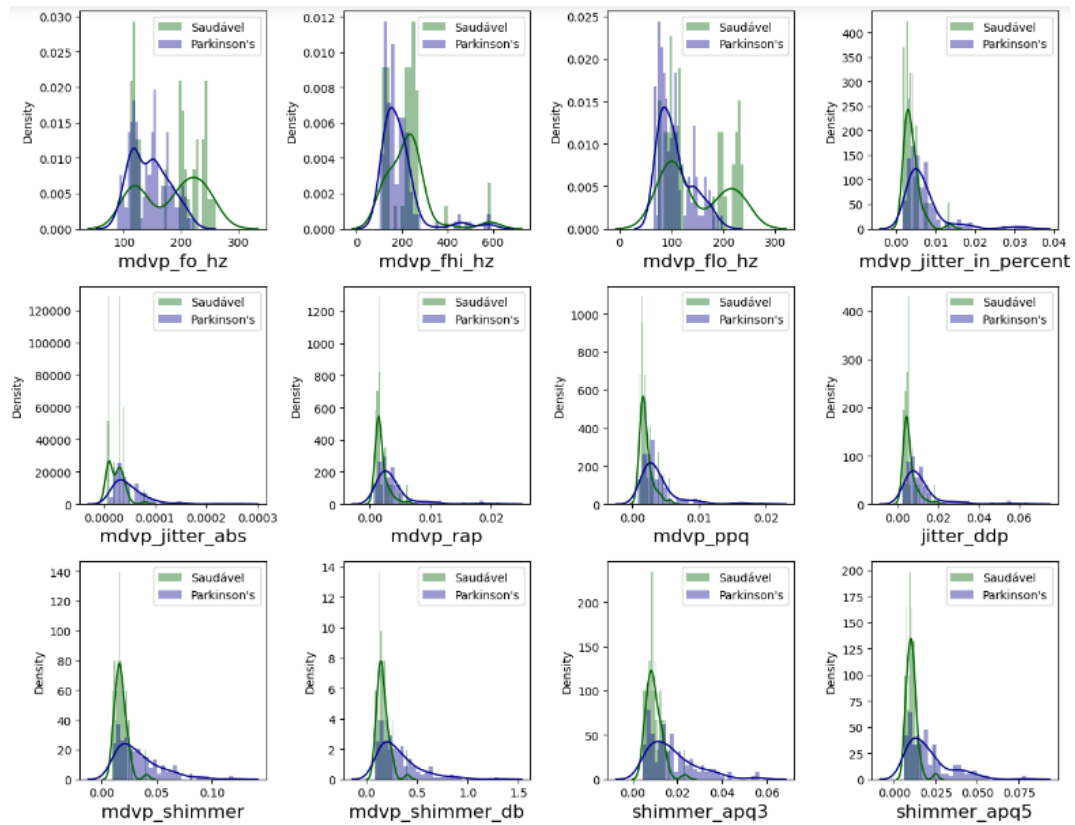


Figura 5 – Primeira Distribuição de *Features* do Primeiro *Dataset*

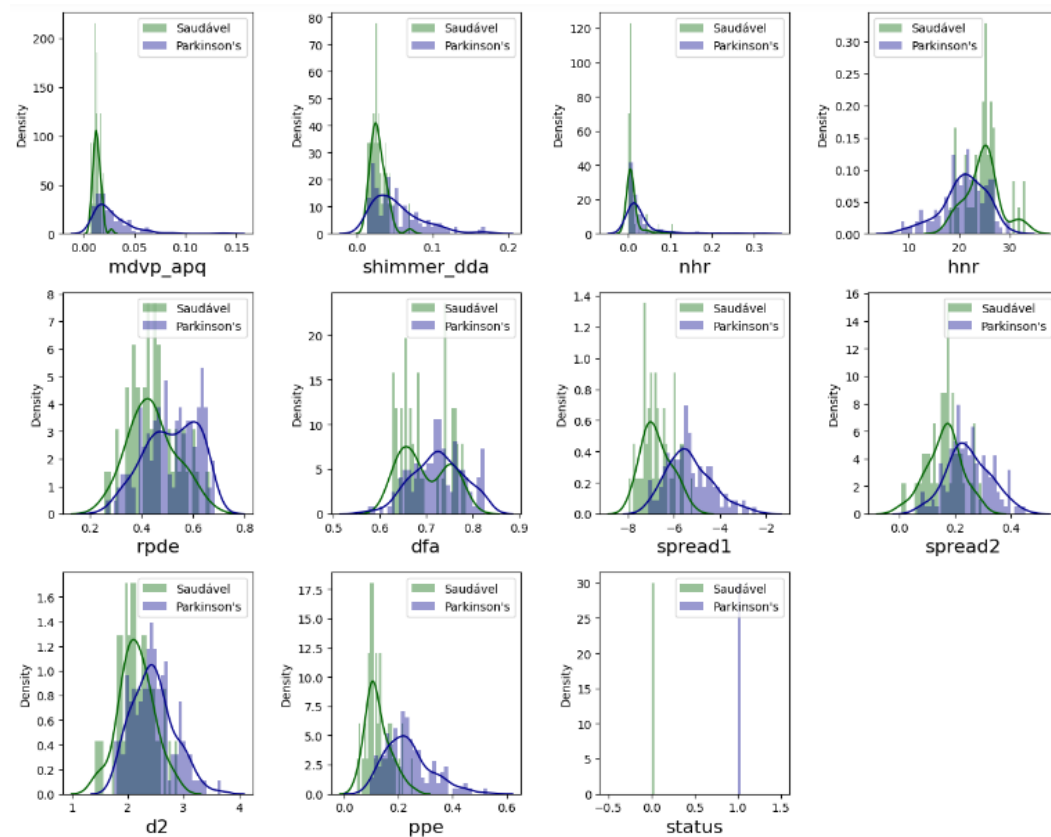


Figura 6 – Segunda Distribuição de *Features* do Primeiro *Dataset*

Ao analisar o gráfico do *scatterplot* de t-SNE, percebemos que no grupo inferior, a separação entre

as classes de indivíduos saudáveis e doentes é mais clara. Entretanto, nos grupos superiores, a distinção não é tão direta quanto parece.

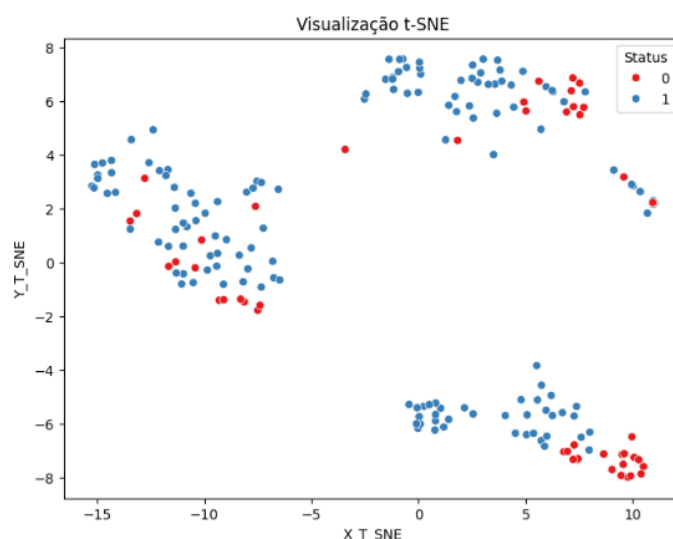


Figura 7 – Análise T-SNE do Primeiro *Dataset*

4.1.2 Segundo Dataset

Da mesma forma que no primeiro conjunto de dados, a proporção entre gravações de áudio de indivíduos doentes e saudáveis permaneceu em 75% para 25%, mesmo diante da significativa ampliação da quantidade de dados disponíveis.

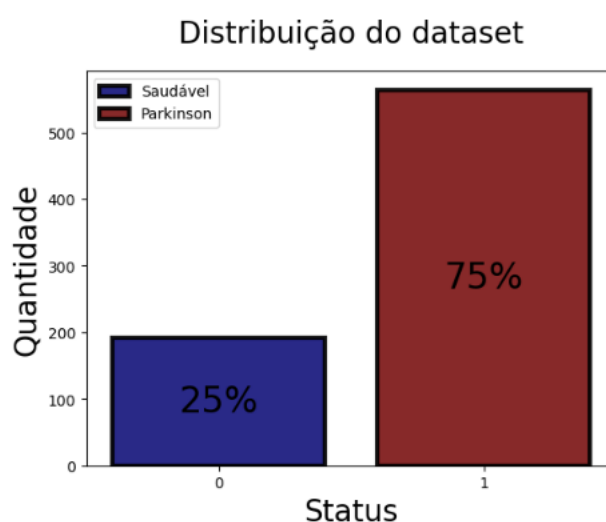


Figura 8 – Distribuição do Segundo *Dataset*

Já na questão do gênero o *dataset* é muito bem equilibrado tendo uma proporção de 52% de mulheres para 48% de gravações de homens, sendo proporcional para verificar se o gênero afeta a doença.

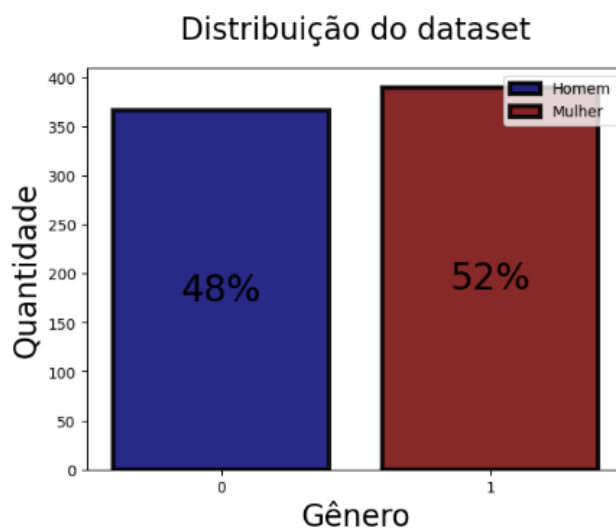


Figura 9 – Distribuição de Gênero Segundo *Dataset*

Nas distribuições conseguimos ver diferenças bem marcantes de quem tem a doença pra quem não tem, como por exemplo no numero de pulsações da onda sonora da voz sofre como podemos ver abaixo, vemos que pessoas saudáveis tem um maior numero de pulsos na média.

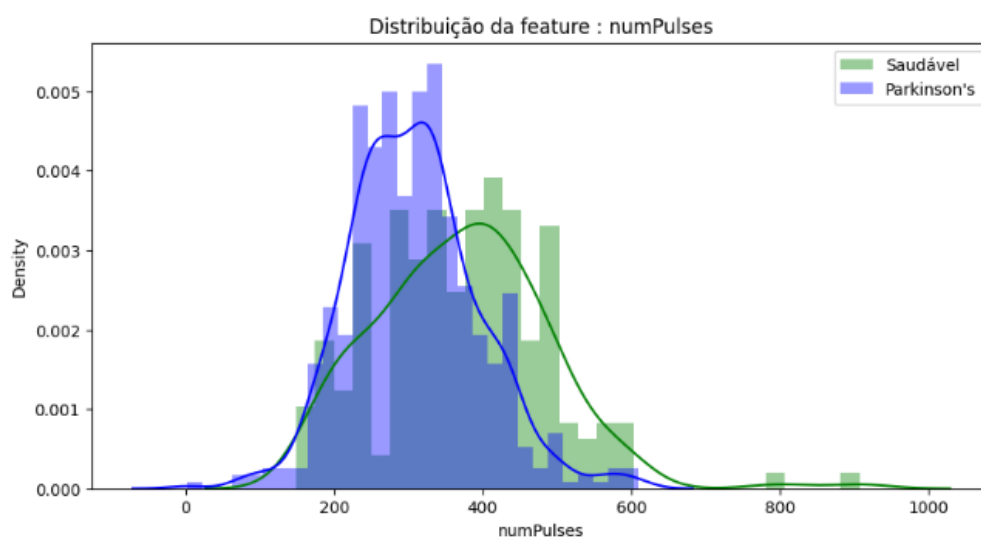


Figura 10 – Distribuição do Número de Pulsações do Segundo *Dataset*

Outra distribuição onde podemos ver a diferença é no periodo medio dos pulsos onde realmente confirmamos que a frequência dos pulsos é maior em pessoas saudáveis.

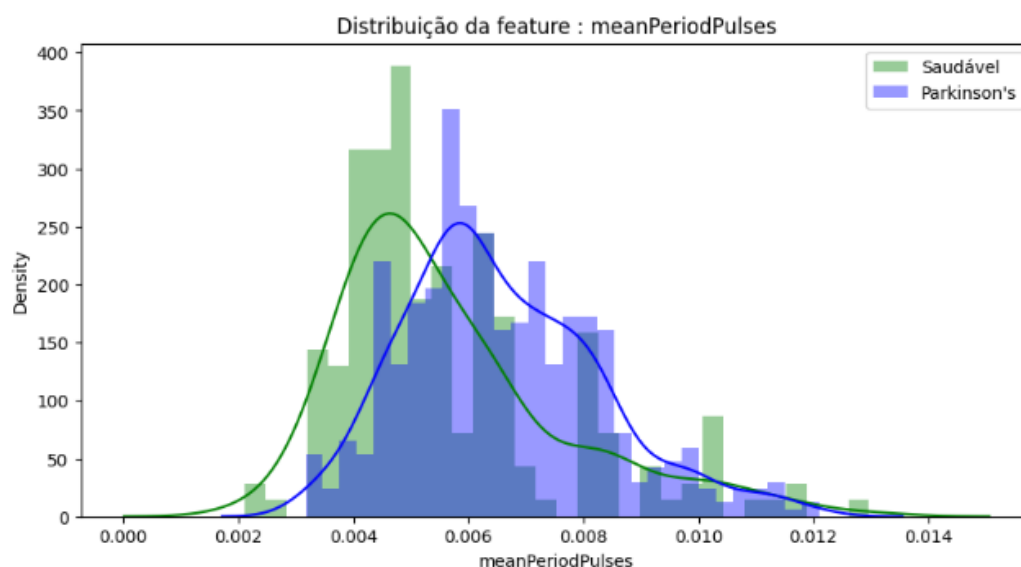


Figura 11 – Distribuição do Número de Pulsações médias do Segundo *Dataset*

4.1.3 Terceiro Dataset

No conjunto de dados em questão, observa-se uma distribuição mais equilibrada entre indivíduos saudáveis e aqueles com Parkinson. Representando 47,1% dos áudios, as gravações de pessoas afetadas pela condição são praticamente equivalentes às de indivíduos sem a doença, totalizando 52,9%. Essa distribuição balanceada reforça a robustez e a representatividade do conjunto de dados, proporcionando uma base sólida para a análise e treinamento do modelo.

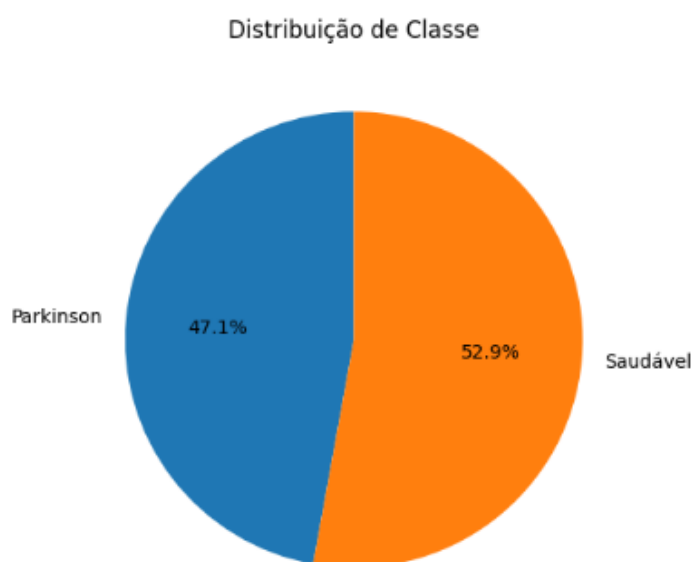


Figura 12 – Distribuição do Terceiro *Dataset*

O conjunto de dados é composto por um total de 73 arquivos de áudio no formato WAV, sendo 37 provenientes de leituras de frases e 36 de diálogos espontâneos. Para otimizar o treinamento e proporcionar um aprendizado eficaz para a inteligência artificial, todos os áudios disponíveis foram empregados. Esta abordagem visa garantir que a IA seja exposta à diversidade de contextos e nuances presentes nos diferentes tipos de gravações, contribuindo para um aprendizado abrangente e preciso.

4.2 SELEÇÃO DE FEATURES

4.2.1 Primeiro Dataset

Para identificar as características que apresentam uma correlação mais significativa com a doença de Parkinson, foi empregado um *heatmap*, permitindo assim visualizar as interações entre as próprias características.

No gráfico abaixo, é possível observar as correlações mais notáveis, evidenciadas pelos tons mais intensos de amarelo e verde. Isso indica que MDVP_Shimmer, MDVP_Shimmer_dB, Shimmer_APQ3, Shimmer_APQ5, MDVP_APQ, e Shimmer_DDA exibem as correlações mais marcantes entre si. Além disso, MDVP_Jitter_in_Percent, MDVP_Jitter_Abs, MDVP_RAP, MDVP_PPQ, e Jitter_DDP também demonstram relações fortes.

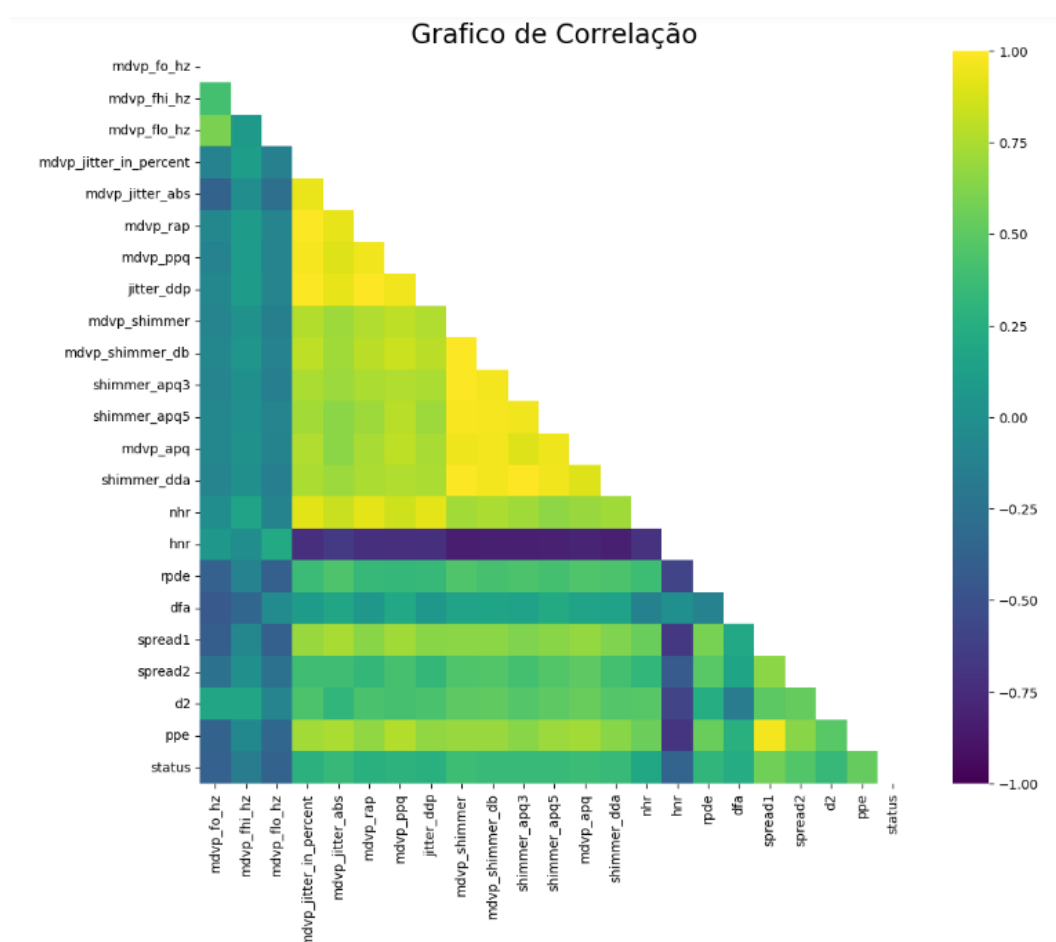


Figura 13 – Gráfico de Correlação do Primeiro Dataset

No que diz respeito às relações com o status, ou seja, se o indivíduo possui a doença ou não, a situação se apresenta da seguinte maneira:

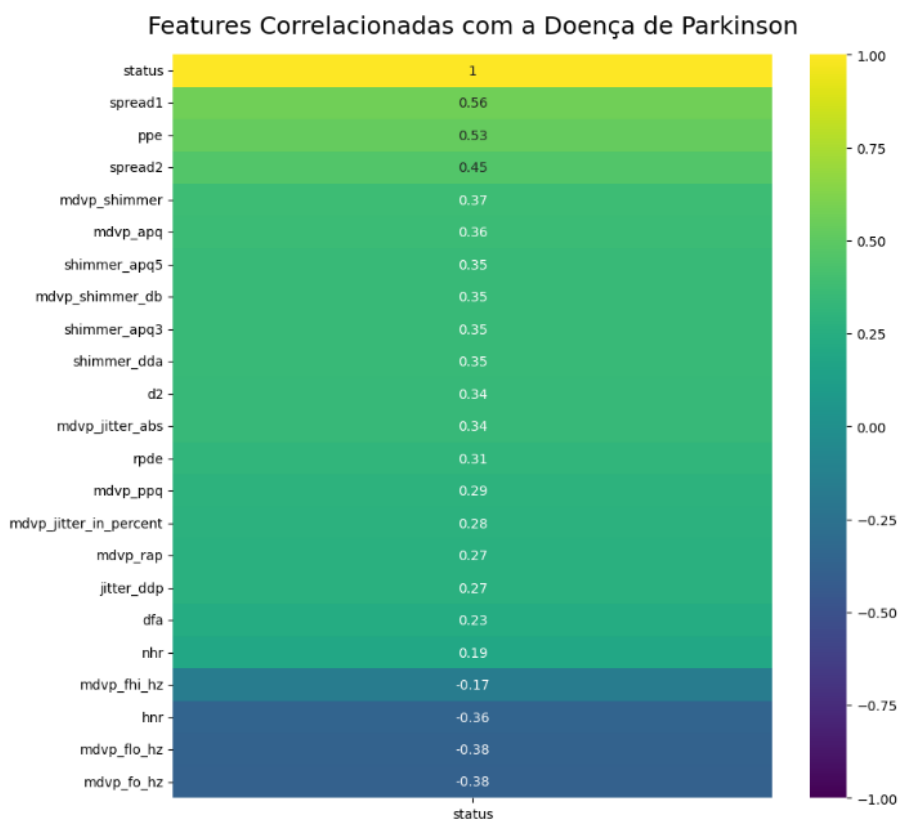


Figura 14 – *Features* com mais Correlações do Primeiro *Dataset*

Isso sugere que spread1, ppe, spread2, MDVP_Shimmer, MDVP_APQ e Shimmer_APQ5 são as características que demonstram as relações mais robustas com a presença da doença.

4.2.2 Segundo Dataset

Por esse *dataset* apresentar muitas *features* um método mais apropriado para reduzir a dimensionalidade foi a utilização de PCA. Para saber quantas *features* selecionar foi analisado o gráfico da variância para saber quando estava alta o bastante.

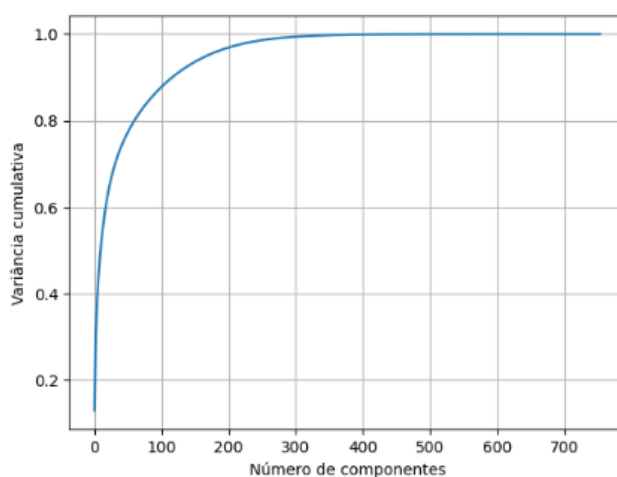


Figura 15 – Vâriança dos Componentes de PCA do Segundo *Dataset*

Pelo gráfico notamos que uma taxa superior a 95% só é alcançada em torno dos 200 *features*, porém

ao fazermos um teste de acurácia com poucas quantidades de features verificamos que conseguimos um bom valor com esses valores.

Componentes	Acurácia
1	78.94736642105263
2	83.55263157894737
3	79.605263157869474
4	83.55263157894737
5	84.210526315786947
6	86.8421052631579
7	83.55263157894737
8	83.55263157894737
9	83.55263157894737
10	83.55263157894737
11	82.89473684210526
12	80.92105263157895
13	83.55263157894737
14	83.55263157894737
15	84.86842105263158
16	86.18421052631578
17	85.52631578947368
18	86.18421052631578
19	86.18421052631578
20	84.210526315786947

Podemos ver que com 6 *features* seleccionadas temos o maior desempenho em acurácia, portanto será a utilizada para reduzir a dimensionalidade do *dataset*.

Feito o PCA podemos ver como se distribuem os dois principais *features* em termos de pessoas doentes e saudáveis.

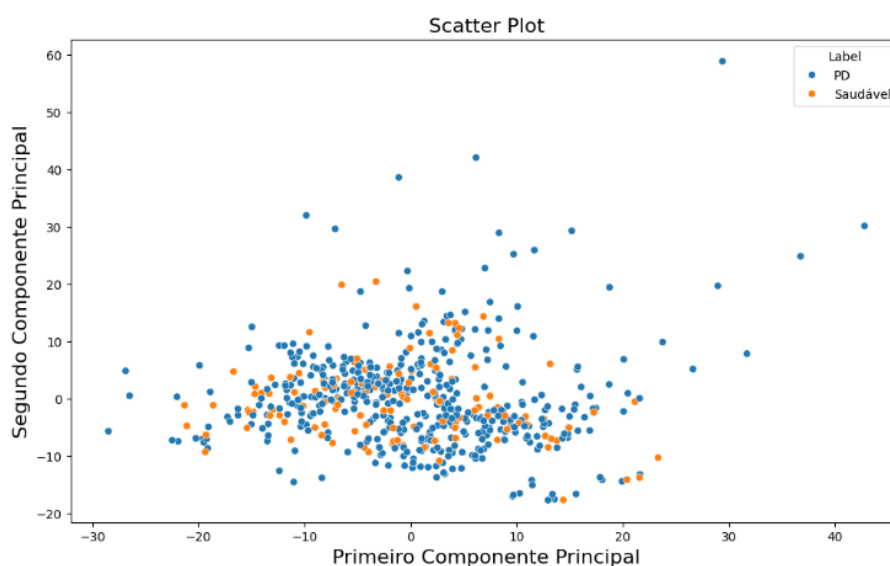


Figura 16 – *Scatterplot* do PCA dos dois principais *features* seleccionados do Segundo *Dataset*

4.2.3 Terceiro Dataset

Para a análise completa e minuciosa dos dados de áudio brutos, optou-se por empregar todas as *features* extraídas, o que inclui os sinais de séries temporais, a Transformada de *Fourier* Rápida (FFT), o banco de filtros mel (FBank) e os coeficientes cepstrais de frequência mel (MFCCs). A escolha de utilizar todas essas *features* foi fundamentada na natureza densa em dados e na riqueza de informações que cada uma delas oferece. Esse conjunto de *features* foi essencial para o treinamento completo do modelo de *deep learning*, permitindo que ele capture e compreenda a complexidade dos padrões presentes nos dados de áudio, resultando em uma análise precisa e robusta.

Podemos visualizar as *features* extraídas dos áudio nos gráficos a seguir.

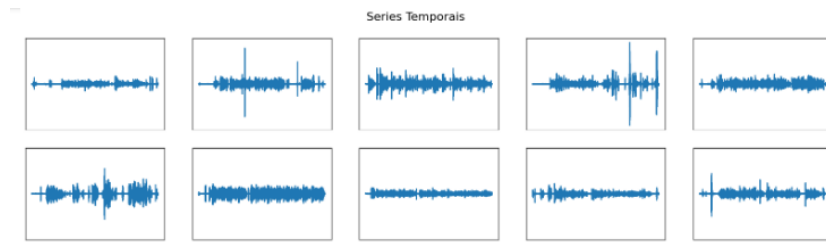


Figura 17 – Dado Temporais



Figura 18 – Transformações Fourier

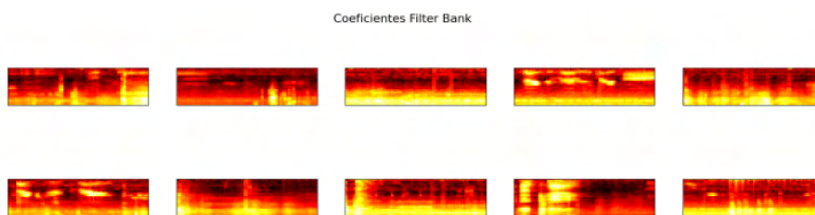


Figura 19 – Coeficientes *Filter Bank*

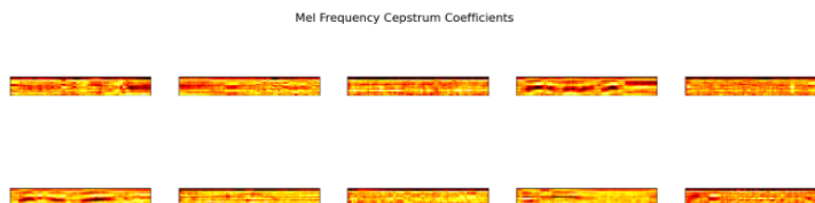


Figura 20 – *Mel Frequency Cepstral Coefficients*

5 TREINAMENTO

Os dados que foram retirados da sessão anterior serão utilizados agora para o treinamento serão utilizados vários modelos com o objetivo de achar o que tem melhor eficácia.

5.1 MÉTRICAS UTILIZADAS

No *dataset* 1, a *feature* 'status' é utilizada para indicar a condição de saúde dos indivíduos. Aqueles com 'status' igual a 0 são considerados saudáveis, enquanto os que têm 'status' igual a 1 são diagnosticados com a doença de Parkinson.

Já no *dataset* 2, a característica relevante é denominada 'classe'. Neste contexto, 'classe' igual a 0 representa indivíduos saudáveis, enquanto 'classe' igual a 1 denota a presença da doença de Parkinson.

Para o *dataset* 3 foi criada uma tabela com os nomes dos arquivos, a coluna label que indica se o paciente tem ou não a doença, sendo essa coluna a relevante, e com o tamanho de cada arquivo.

ID00_hc_0_0_0.wav	Saudável	102.430937
ID01_hc_0_0_0.wav	Saudável	97.260063
ID02_pd_2_0_0.wav	Parkinson	139.108812
ID03_hc_0_0_0.wav	Saudável	132.645438
ID04_pd_2_0_1.wav	Parkinson	112.720562
...
ID68_hc_0_1_1.wav	Saudável	176.083375
ID69_pd_3_1_1.wav	Parkinson	72.698813
ID70_pd_3_2_2.wav	Parkinson	90.533750
ID71_pd_2_0_0.wav	Parkinson	119.982750
ID72_hc_0_0_0.wav	Saudável	83.283312

73 rows x 2 columns

Figura 21 – Tabela criada

No campo da aprendizagem de máquina, a avaliação do desempenho de um modelo é de crucial importância para garantir sua eficácia na tarefa designada. Entre as métricas fundamentais para esta avaliação, destacam-se a *Confusion Matrix* (Matriz de Confusão), *Accuracy* (Acurácia), *Precision* (Precisão) e *Recall* (Revocação) e *F1-Score*. Essas métricas fornecem uma visão abrangente das capacidades do modelo, permitindo uma análise detalhada das previsões e classificações. Ao entender o significado e a aplicação dessas métricas, somos capazes de aprimorar a eficiência e a confiabilidade dos modelos de aprendizado de máquina, sendo essenciais para aplicações que demandam precisão e confiabilidade na identificação de diferentes classes ou estados, como no diagnóstico de condições médicas, detecção de anomalias em sistemas, entre outras aplicações críticas.(JAYASWAL, 2020)

- *Confusion Matrix* (Matriz de Confusão): A matriz de confusão é uma tabela que descreve o desempenho de um modelo de classificação. Ela apresenta a contagem de verdadeiros positivos,

verdadeiros negativos, falsos positivos e falsos negativos. É uma ferramenta valiosa para entender como o modelo está classificando os diferentes grupos.

- *Accuracy* (Acurácia): A acurácia é uma medida que indica a proporção de predições corretas que o modelo faz em relação ao total de predições. Matematicamente, é calculada como $(TP + TN) / (TP + TN + FP + FN)$, onde TP são os verdadeiros positivos, TN são os verdadeiros negativos, FP são os falsos positivos e FN são os falsos negativos.
- *Precision* (Precisão): A precisão é a proporção de verdadeiros positivos em relação ao total de predições positivas feitas pelo modelo. Ela é calculada como $TP / (TP + FP)$. A precisão é especialmente importante quando queremos ter certeza de que as predições positivas são confiáveis.
- *Recall* (Revocação): A revocação é a proporção de verdadeiros positivos em relação ao total de casos verdadeiramente positivos na amostra. Ela é calculada como $TP / (TP + FN)$. A revocação é crucial quando a identificação de todos os casos positivos é uma prioridade.
- *F1-Score*: O *F1-Score* é uma métrica que combina precisão (*precision*) e revocação (*recall*) em um único valor. Ele fornece uma medida mais equilibrada entre essas duas métricas e é particularmente útil quando há um desequilíbrio entre as classes no conjunto de dados. Ele é calculado como $2 * (Precision * Recall) / (Precision + Recall)$. Essa métrica é especialmente valiosa em cenários onde tanto falsos positivos quanto falsos negativos têm impactos significativos.

5.2 SAMPLING DOS DADOS

Para o *sampling* dos dados será utilizado o *k-fold cross validation*, separando 20% dos dados para o teste e 80% para o treino, e distribuindo em 10 *folds*, ou subgrupos.

O uso de *k-fold cross validation* é uma técnica importante em aprendizado de máquina e estatística para avaliar o desempenho de um modelo.

Ao dividir o conjunto de dados em *k* subconjuntos, podemos aproveitar ao máximo nossos dados para treinamento e teste. Cada observação é usada para treinamento e teste exatamente uma vez. O *k-fold* oferece uma avaliação mais robusta do desempenho do modelo, uma vez que o modelo é avaliado em várias partições diferentes dos dados. Isso ajuda a evitar resultados enviesados que poderiam ocorrer se usássemos apenas um único conjunto de treinamento e teste.

Como o modelo é testado em várias partições diferentes dos dados, a estimativa do desempenho do modelo é menos sensível à forma como os dados são divididos inicialmente. Isso reduz a variação nos resultados e oferece uma avaliação mais estável do desempenho. Além disso pode ajudar a detectar se o modelo está sofrendo de *overfitting*. Se o modelo tem um desempenho significativamente pior nos dados de teste em comparação com os dados de treinamento, isso pode indicar *overfitting*.

Em conjuntos de dados pequenos, é particularmente importante usar validação cruzada para garantir que possamos avaliar o desempenho do modelo de forma robusta, mesmo com uma quantidade limitada de dados disponíveis.

5.3 TREINAMENTO DOS MODELOS

Com base na análise dos dados e utilizando técnicas avançadas de aprendizado de máquina, foi desenvolvido e treinado uma variedade de modelos para abordar a questão da detecção de Parkinson. Cada modelo foi cuidadosamente ajustado e validado para garantir sua eficácia e precisão na identificação desta condição neurodegenerativa.

Nesse capítulo vamos explorar os resultados obtidos por meio desses modelos e discutir como cada um deles contribui para o nosso entendimento e capacidade de diagnóstico do Parkinson. Examinando os dados utilizados para treinamento e detalhando os métodos de treinamento e validação empregados para garantir a robustez e confiabilidade dos modelos.

Será apresentado suas métricas de desempenho, como acurácia, precisão e *recall*, que ajudarão na prática para fornecer suporte valioso na identificação desta condição.

5.3.1 Regressão Logística

5.3.1.1 Primeiro Dataset

Aplicando a regressão logística no primeiro *dataset* verificamos que possui uma acurácia de treino de 83,97%, um score de *cross validation* de 82,63%, uma acurácia de teste de 92,31%, uma *precision* de 91,43%, uma *recall* de 100% e um F1-score de 95,52%, indicando uma boa taxa de acertos, e até uma ausência de falsos negativos.

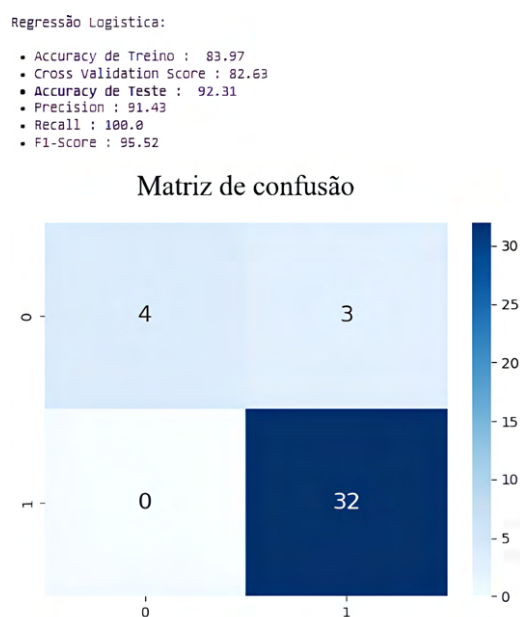


Figura 22 – Regressão Linear no Primeiro *Dataset*

5.3.1.2 Segundo Dataset

Já no segundo *dataset*, obtivemos uma acurácia de treino de 81,62%, um *score* de *cross validation* de 80,78%, uma acurácia de teste de 86,84%, uma *precision* de 86,72%, uma *recall* de 97,37% e um F1-score de 91,74%, sendo taxas aceitavelmente altas também

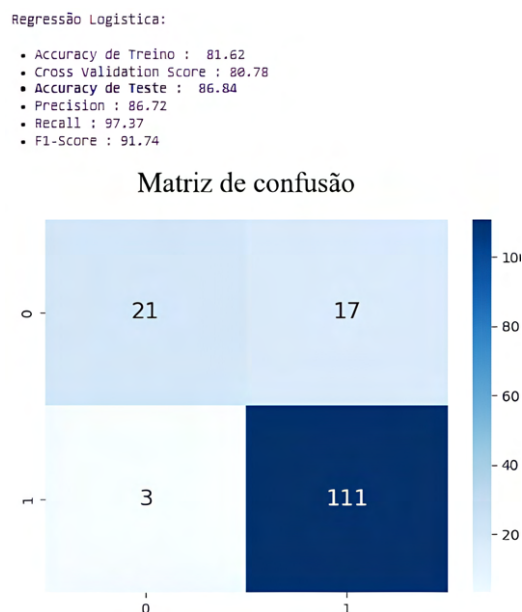


Figura 23 – Regressão Linear no Segundo *Dataset*

5.3.2 Support Vector Machine

5.3.2.1 Primeiro Dataset

Agora para o *Support Vector Machine* no primeiro *dataset* conseguimos uma acurácia de treino de 83,97%, um *score* de *cross validation* de 83.21%, uma acurácia de teste de 92,31%, uma *precision* de 91.43%, uma *recall* de 100% e um F1-score de 95,52%, dados parecidos com a regressão linear.

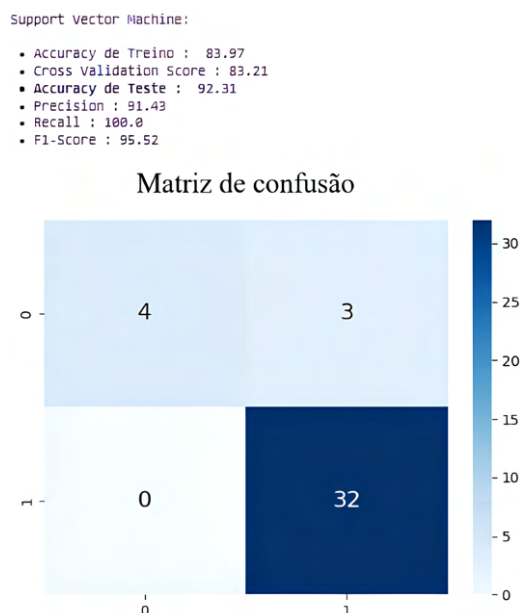


Figura 24 – *Support Vector Machine* no Primeiro *Dataset*

5.3.2.2 Segundo Dataset

E no segundo conseguimos uma acurácia de treino de 82.45%, um *score* de *cross validation* de 80.78%, uma acurácia de teste de 84.21%, uma *precision* de 84.09%, uma *recall* de 97.37% e um

F1-score de 90.24%, dados bons mas com menos acurácia que a regressão logística.

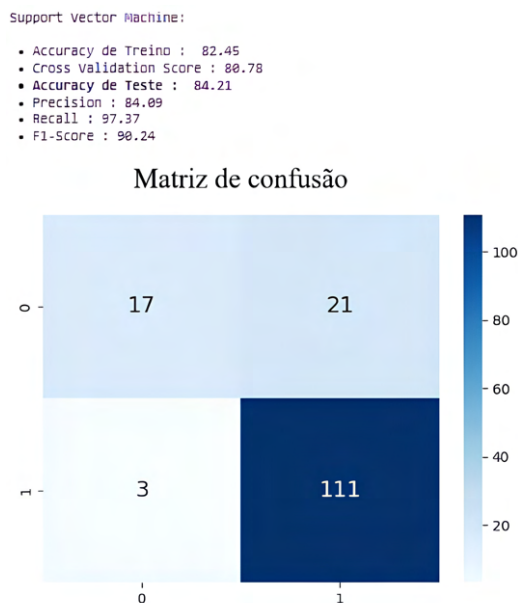


Figura 25 – *Support Vector Machine* no Segundo Dataset

5.3.3 K-Neighbors

5.3.3.1 Primeiro Dataset

Com o *K-Neighbors* no primeiro *dataset* temos uma acurácia de treino de 83,33%, um *score* de *cross validation* de 76.21%, uma acurácia de teste de 89,74%, uma *precision* de 91.18%, uma *recall* de 96,88% e um F1-score de 93,94%, uma leve diferença para os anteriores, diminuindo um pouco a acurácia do treino.

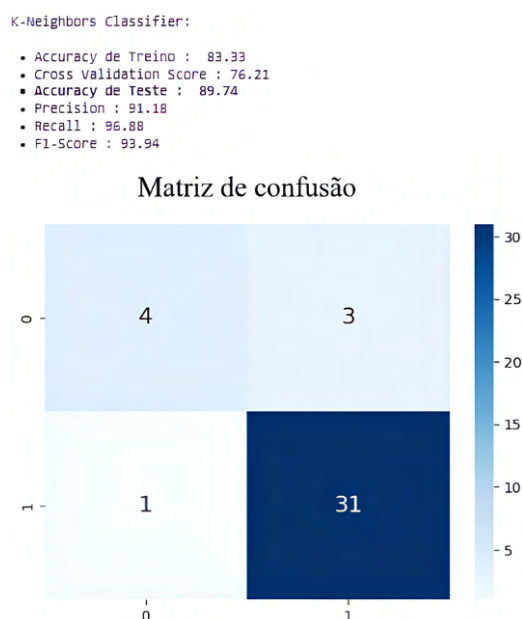


Figura 26 – *K-Neighbors* no Primeiro Dataset

5.3.3.2 Segundo Dataset

Para o segundo a acurácia de treino foi de 89.07%, um *score* de *cross validation* de 81.44%, uma acurácia de teste de 82.24%, uma *precision* de 87.83%, uma *recall* de 88.60% e um F1-score de 88.21%, observamos mais uma diminuição da acurácia.

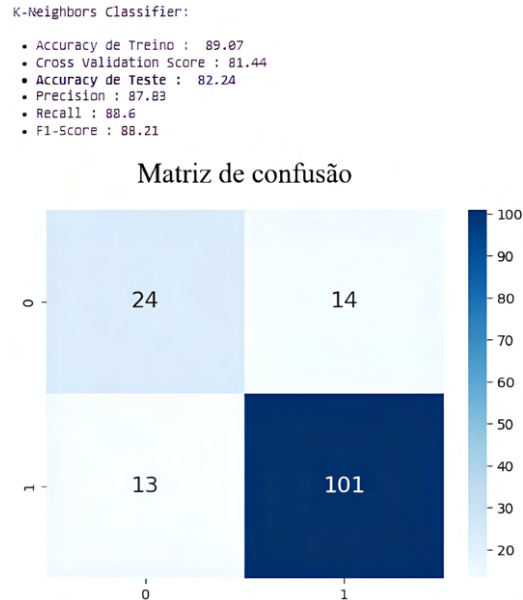


Figura 27 – K-Neighbors no Segundo Dataset

5.3.4 Random Forest

5.3.4.1 Primeiro Dataset

Utilizando Random Forest no primeiro *dataset* obtemos uma acurácia de treino de 100,00%, um *score* de *cross validation* de 83,92%, uma acurácia de teste de 92,31%, uma *precision* de 91.43%, uma *recall* de 100% e um F1-score de 95,52%, mais uma vez dados parecidos com os anteriores mantendo taxas altas de acerto.

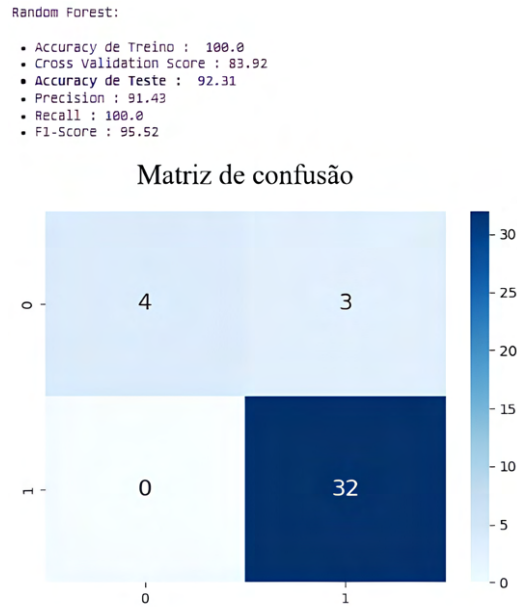


Figura 28 – *Random Forest* no Primeiro Dataset

5.3.4.2 Segundo Dataset

No segundo *dataset* conseguimos a acurácia de treino de 100.00%, um *score* de *cross validation* de 81.97%, uma acurácia de teste de 82.89%, uma *precision* de 86.67%, uma *recall* de 91.23% e um *F1-score* de 88.89%, mantendo menores taxas que a regressão logística.

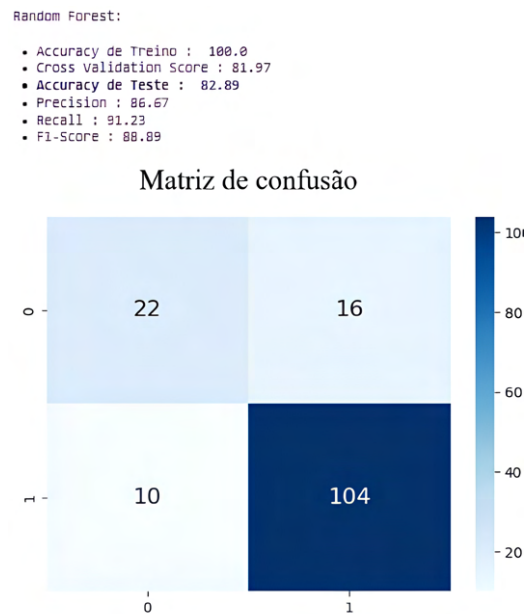


Figura 29 – *Random Forest* no Segundo Dataset

5.3.5 Decision Tree

5.3.5.1 Primeiro Dataset

A *Decision Tree* para o primeiro alcançou uma acurácia de treino de 100,00%, um *score* de *cross validation* de 82.62%, uma acurácia de teste de 89,74%, uma *precision* de 91.18%, uma *recall* de

96,88% e um F1-score de 93,94%, dados parecidos com *k-neighbors*.



Figura 30 – *Decision Tree* no Primeiro Dataset

5.3.5.2 Segundo Dataset

O segundo alcançou a acurácia de treino de 100.00%, um *score* de *cross validation* de 75.83%, uma acurácia de teste de 79.61%, uma *precision* de 86.73%, uma *recall* de 85.96% e um F1-score de 86.34%, podemos ver taxas baixas de acurácia.

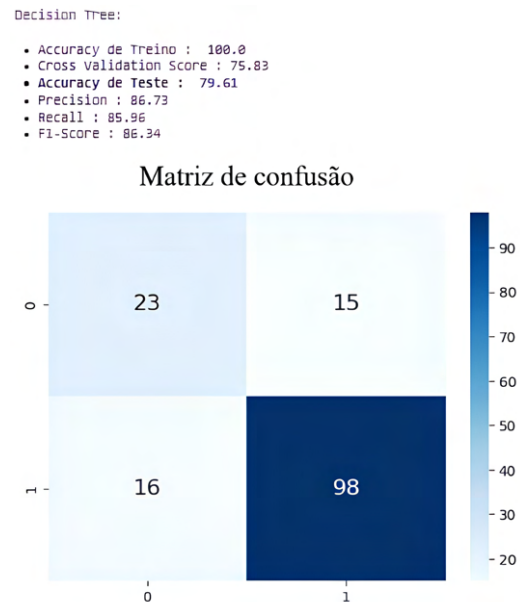


Figura 31 – *Decision Tree* no Segundo Dataset

5.3.6 Naive Bayes

5.3.6.1 Primeiro Dataset

O *Naive Bayes* no primeiro teve uma acurácia de treino de 73.08%, um *score* de *cross validation* de 73.71%, uma acurácia de teste de 87.18%, uma *precision* de 96.55%, uma *recall* de 87.50% e um *F1-score* de 91.80%, os dados que mais se diferenciam dos outros porem uma diferença baixa na acurácia, apresentando a menor taxa para ela.

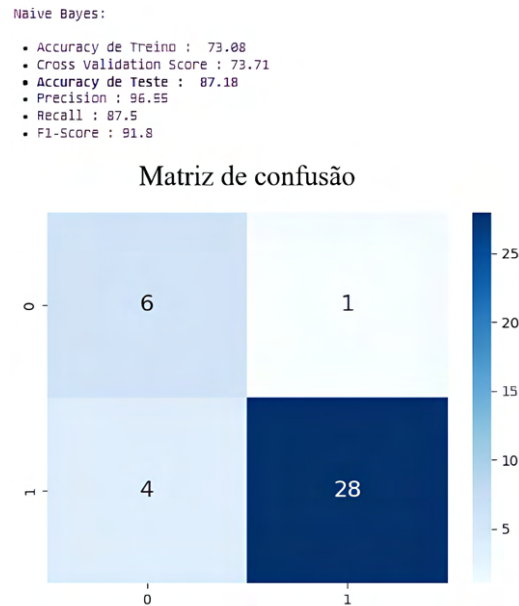


Figura 32 – *Naive Bayes* no Primeiro Dataset

5.3.6.2 Segundo Dataset

O segundo teve a acurácia de treino de 74.83%, um *score* de *cross validation* de 75.15%, uma acurácia de teste de 75.00%, uma *precision* de 77.94%, uma *recall* de 92.98% e um *F1-score* de 84.80%, obtendo taxas muito baixas em tudo, as piores no segundo *dataset*.

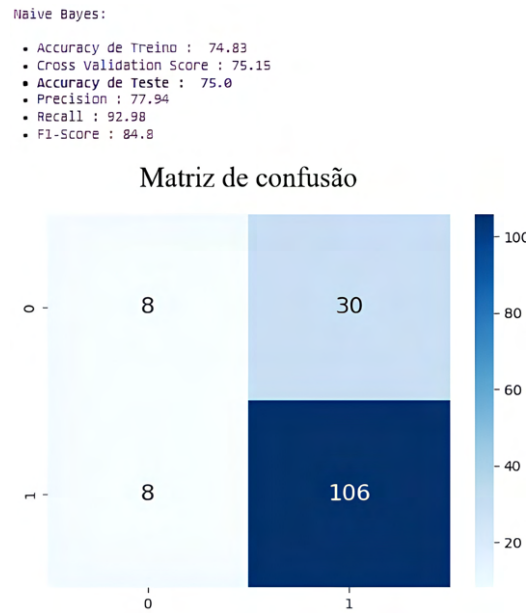


Figura 33 – Naive Bayes no Segundo Dataset

5.3.7 XGBoost

5.3.7.1 Primeiro Dataset

E por último foi utilizado *XGBoost* no primeiro *dataset* para obter uma acurácia de treino de 100.00%, um *score* de *cross validation* de 81.33%, uma acurácia de teste de 92.31%, uma *precision* de 91.43%, uma *recall* de 100.00% e um *F1-score* de 95.52%, repetindo os bons dados de alguns outros modelos.

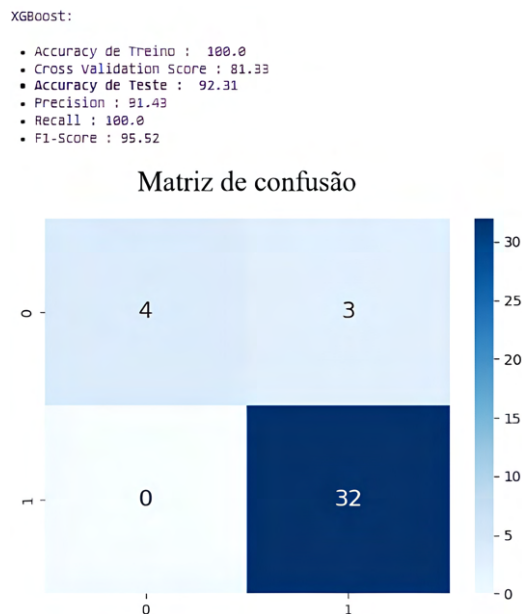


Figura 34 – XGBoost no Primeiro Dataset

5.3.7.2 Segundo Dataset

O *XGBoost* no segundo teve a acurácia de treino de 94.54%, um *score* de *cross validation* de 79.64%, uma acurácia de teste de 83.55%, uma *precision* de 86.18%, uma *recall* de 92.98% e um

F1-score de 89.45%, com uma boa acurácia mas não a melhor.

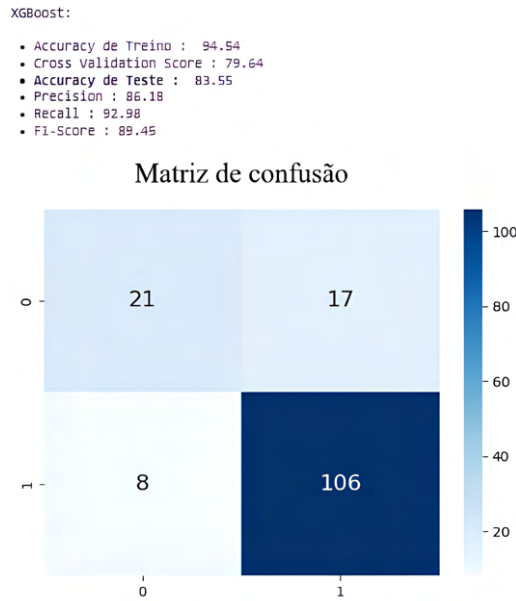


Figura 35 – XGBoost no Segundo Dataset

5.3.8 Deep Learning no Terceiro Dataset

No caso do terceiro conjunto de dados, adotou-se uma abordagem baseada em *deep learning*. Para este fim, empregou-se uma combinação de Redes Neurais Convolucionais (CNN) e Redes Neurais Recorrentes (RNN) visando treinar o modelo e alcançar os resultados mais promissores.

5.3.8.1 CNN

Primeiramente treinamos utilizando CNN com dez épocas, o resultado foi de uma acurácia de treino 0.8580, perda de treino de 0.3144, acurácia de validação de 0.8669 e perda de validação de 0.2987.

```
Epoch 1/10
5063/5063 [=====] - 90s 18ms/step - loss: 0.5927 - acc: 0.6732 - val_loss: 0.5599 - val_acc: 0.7083
Epoch 2/10
5063/5063 [=====] - 87s 17ms/step - loss: 0.4934 - acc: 0.7553 - val_loss: 0.4639 - val_acc: 0.7718
Epoch 3/10
5063/5063 [=====] - 88s 17ms/step - loss: 0.4428 - acc: 0.7863 - val_loss: 0.4004 - val_acc: 0.8075
Epoch 4/10
5063/5063 [=====] - 90s 18ms/step - loss: 0.4113 - acc: 0.8055 - val_loss: 0.3798 - val_acc: 0.8212
Epoch 5/10
5063/5063 [=====] - 90s 18ms/step - loss: 0.3880 - acc: 0.8182 - val_loss: 0.3706 - val_acc: 0.8301
Epoch 6/10
5063/5063 [=====] - 88s 17ms/step - loss: 0.3695 - acc: 0.8284 - val_loss: 0.3489 - val_acc: 0.8383
Epoch 7/10
5063/5063 [=====] - 89s 18ms/step - loss: 0.3514 - acc: 0.8374 - val_loss: 0.3298 - val_acc: 0.8494
Epoch 8/10
5063/5063 [=====] - 87s 17ms/step - loss: 0.3378 - acc: 0.8448 - val_loss: 0.3211 - val_acc: 0.8529
Epoch 9/10
5063/5063 [=====] - 91s 18ms/step - loss: 0.3246 - acc: 0.8519 - val_loss: 0.3123 - val_acc: 0.8593
Epoch 10/10
5063/5063 [=====] - 91s 18ms/step - loss: 0.3144 - acc: 0.8580 - val_loss: 0.2987 - val_acc: 0.8669
```

Figura 36 – Treinamento CNN

Podemos ver que a validação segue melhor que o treino até as últimas épocas, demonstrando que com mais épocas as duas métricas andam lado a lado.

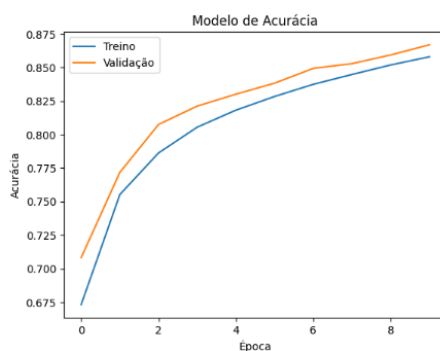


Figura 37 – Acurácia do Treinamento CNN

A perda assim como a acurácia segue melhorando até as épocas finais.

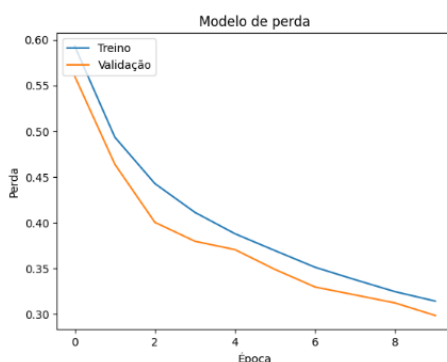


Figura 38 – Perda do Treinamento CNN

Plotando a matriz de confusão obtemos a precisão de 86.83%, um *Recall* de 88.20% e *F1-Score* 87.51%.

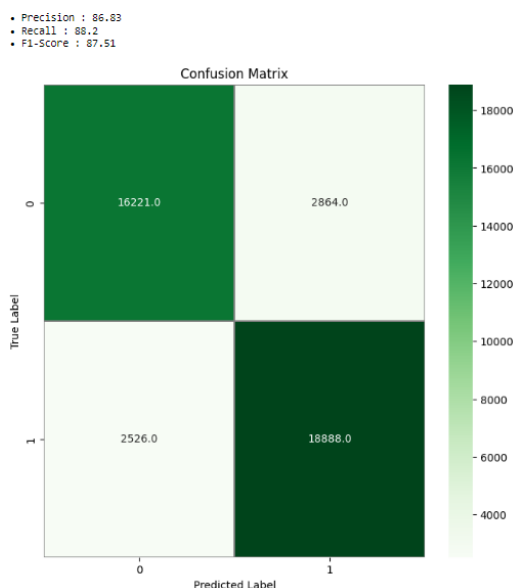


Figura 39 – Matriz de Confusão CNN

5.3.8.2 RNN

Os resultados do treinamento com Redes Neurais Recorrentes (RNN) foram os seguintes: uma acurácia de treino de 0.7806 e uma perda de treino de 0.4461. Em relação à validação, obtemos uma

acurácia de 0.7884 e uma perda de validação de 0.4325.

```
Epoch 1/10
5063/5063 [=====] - 58s 11ms/step - loss: 0.6590 - acc: 0.6019 - val_loss: 0.6468 - val_acc: 0.6063
Epoch 2/10
5063/5063 [=====] - 54s 11ms/step - loss: 0.6206 - acc: 0.6465 - val_loss: 0.6020 - val_acc: 0.6727
Epoch 3/10
5063/5063 [=====] - 54s 11ms/step - loss: 0.5878 - acc: 0.6754 - val_loss: 0.5550 - val_acc: 0.6977
Epoch 4/10
5063/5063 [=====] - 56s 11ms/step - loss: 0.5579 - acc: 0.6992 - val_loss: 0.5261 - val_acc: 0.7215
Epoch 5/10
5063/5063 [=====] - 55s 11ms/step - loss: 0.5348 - acc: 0.7171 - val_loss: 0.5223 - val_acc: 0.7261
Epoch 6/10
5063/5063 [=====] - 56s 11ms/step - loss: 0.5146 - acc: 0.7328 - val_loss: 0.4969 - val_acc: 0.7424
Epoch 7/10
5063/5063 [=====] - 55s 11ms/step - loss: 0.4963 - acc: 0.7463 - val_loss: 0.4750 - val_acc: 0.7625
Epoch 8/10
5063/5063 [=====] - 56s 11ms/step - loss: 0.4772 - acc: 0.7600 - val_loss: 0.4744 - val_acc: 0.7632
Epoch 9/10
5063/5063 [=====] - 58s 11ms/step - loss: 0.4623 - acc: 0.7700 - val_loss: 0.4574 - val_acc: 0.7693
Epoch 10/10
5063/5063 [=====] - 54s 11ms/step - loss: 0.4461 - acc: 0.7806 - val_loss: 0.4325 - val_acc: 0.7884
```

Figura 40 – Treinamento RNN

Ao analisarmos o progresso do treinamento ao longo das épocas, observamos uma tendência interessante de que em certos pontos, as métricas de treinamento e validação se aproximam muito, quase se sobrepondo no gráfico. Isso sugere que, ao estender o treinamento por mais épocas, é provável que a proporção entre as métricas se mantenha. Essa tendência pode ser um indicativo de estabilidade no processo de aprendizado do modelo.

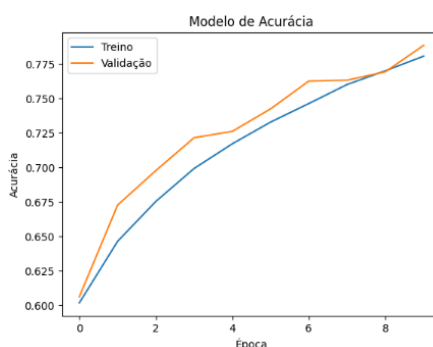


Figura 41 – Acurácia do Treinamento RNN

A análise da evolução da perda ao longo das épocas revela um padrão promissor, ela diminui de forma consistente com o avanço das épocas. Isso sugere que ao estender o treinamento por mais épocas, é provável que a perda continue a reduzir-se, indicando um refinamento contínuo do modelo.

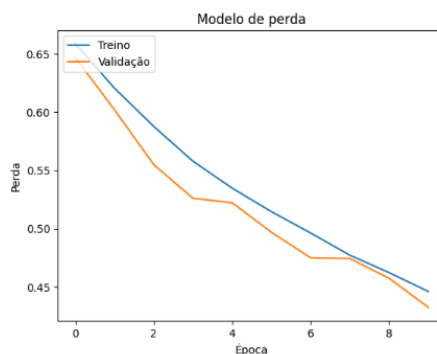


Figura 42 – Perda do Treinamento RNN

A matriz de confusão revela uma precisão de 81.19%, um *Recall* de 78.19% e um *F1-Score* de 79.66%.

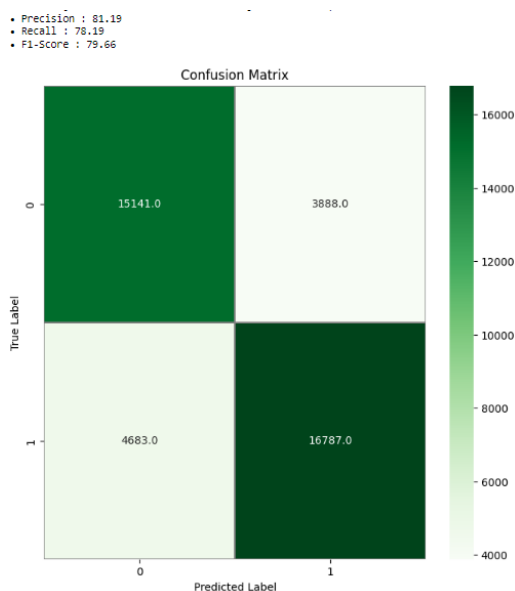


Figura 43 – Matriz de Confusão RNN

5.4 RESULTADOS

Ao analisarmos os resultados dos treinamentos dos modelos, obtemos o seguinte para o primeiro dataset:

Modelos	Acurácia
Regressão Logística	92.3076923076923
Support Vector Machine	92.3076923076923
K-Neighbors Classifier	89.74358974358975
Random Forest	92.3076923076923
Decision Tree	89.74358974358975
Naive Bayes	87.17948717948716
XGBoost	92.3076923076923

Vemos muita repetição de resultados possivelmente pelo *dataset* ser aproximadamente linearmente separável, já que mesmo os resultados diferentes pouco diferem entre si, também vemos que os melhores modelos para os treinamentos são Regressão Logística, *Support Vector Machine*, *Random Forest* e *XGBoost* com 92,31%.

Para o segundo *dataset* os resultados são o seguinte:

Modelos	Acurácia
Regressão Logística	86.8421052631579
Support Vector Machine	84.21052631578947
K-Neighbors Classifier	82.23684210526315
Random Forest	82.23684210526315
Decision Tree	79.60526315789474
Naive Bayes	75.00000000000000
XGBoost	83.55263157894737

Conseguimos diferentes resultados, apesar de não tanta distância entre os valores temos alguns performando bem abaixo da média como *Naive Bayes* e *Decision Tree*. O melhor resultado foi obtido com a Regressão Logística com 86,84%.

Analisando o conjunto Regressão linear foi o que obteve os melhores resultados no geral, provavelmente por ser muito eficaz em treinos binários, já que a classificação tem que retornar se é ou não doente, esse modelo realmente tem seu destaque.

Já no *deep learning*, observou-se um desempenho superior com a rede neural convolucional, alcançando uma taxa de acerto de 81.53%, em comparação com a rede neural recorrente que obteve 76.70%. Esse é um resultado inesperado, visto que, espera-se que redes neurais recorrentes se saiam melhor em tarefas envolvendo áudios e séries temporais. Isso ressalta a importância de explorar diferentes arquiteturas de redes neurais para obter os melhores resultados em cada cenário.

Modelos	Acurácia
Convolutional Neural Network	81.53
Recurrent Neural Network	76.70

Os resultados obtidos demonstram a eficácia do *deep learning*, mesmo com um número reduzido de épocas de treinamento, na tarefa de classificação de indivíduos com a condição e aqueles sem ela a partir de áudios brutos.

6 CONCLUSÃO

Ao analisar e explorar a aplicação de métodos de *machine learning* para a detecção de Parkinson, foi evidente o poder e a versatilidade dessas técnicas no contexto médico. Ao treinar e avaliar diversos modelos em diferentes conjuntos de dados, a regressão logística se apresentou como uma ferramenta notável em nossa análise.

Os resultados obtidos ao aplicar a regressão logística e outros modelos de treino nos datasets analisados retornaram resultados satisfatórios. A precisão e o *recall* alcançados demonstraram uma capacidade interessante de discriminação entre os casos de Parkinson e os indivíduos saudáveis.

Também ficou destacado a importância de uma abordagem holística na análise de dados médicos. A combinação de *features* específicas extraídas dos *datasets* e a escolha criteriosa do algoritmo de *machine learning* desempenhou um papel fundamental nos resultados obtidos.

No entanto, é importante lembrar que cada caso clínico é único, e a aplicação dessas técnicas e a análise delas devem ser sempre feita com cuidado e sob a supervisão de profissionais de saúde qualificados.

Ficou reforçado a promessa e o potencial dos métodos de *machine learning* na detecção precoce de condições médicas em especial a doença de Parkinson, abrindo portas para futuras pesquisas e aplicações práticas que beneficiarão a saúde e o bem-estar da comunidade.

REFERÊNCIAS

- BISWAS, D. **Parkinson's Disease (PD) classification**. 2019. Disponível em: <<https://www.kaggle.com/datasets/dipayanbiswas/parkinsons-disease-speech-signal-features>>. Acesso em: 19/08/2023.
- DANTAS, S. **Pay attention - Explicando o mecanismo de atenção**. 2019. Disponível em: <<https://lamfo-unb.github.io/2019/05/01/Pay-attention-Explicando-o-mecanismo-de-Atencao/>>. Acesso em: 19/08/2023.
- DEVELOPERS scikit-learn. **Decision Trees**. 2023. Disponível em: <<https://scikit-learn.org/stable/modules/tree.html>>. Acesso em: 25/08/2023.
- DEVELOPERS scikit-learn. **Support Vector Machines**. 2023. Disponível em: <<https://scikit-learn.org/stable/modules/svm.html>>. Acesso em: 21/08/2023.
- DEVELOPERS xgboost. **XGBoost Documentation**. 2022. Disponível em: <<https://xgboost.readthedocs.io/en/stable/>>. Acesso em: 25/08/2023.
- FONOFF, D. E. **O que você precisa saber sobre a doença de Parkinson**. 2020. Disponível em: <<https://www.erichfonoff.com.br/doenca-de-parkinson/>>. Acesso em: 19/08/2023.
- GEEKSFORGEEKS. **K-Nearest Neighbor(KNN) Algorithm**. 2023. Disponível em: <<https://www.geeksforgeeks.org/k-nearest-neighbours/>>. Acesso em: 25/08/2023.
- GOVINDU, S. P. A. **Early detection of Parkinson's disease using machine learning**. *Procedia Computer Science*, v. 218, p. 249–261, 2023. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050923000078?via%3Dihub>>. Acesso em: 18/07/2023.
- GUSMÃO, A. **O que é Naive Bayes e como funciona esse algoritmo de classificação**. 2023. Disponível em: <<https://rockcontent.com/br/blog/naive-bayes/>>. Acesso em: 25/08/2023.
- IBM. **O que são redes neurais convolucionais?** 2023. Disponível em: <<https://www.ibm.com/br-pt/topics/convolutional-neural-networks>>. Acesso em: 19/08/2023.
- IBM. **What is random forest?** 2023. Disponível em: <<https://www.ibm.com/topics/random-forest#:~:text=Random%20forest%20is%20a%20commonly,both%20classification%20and%20regression%20problems.>>. Acesso em: 21/08/2023.
- IYODA, E. M. **INTELIGÊNCIA COMPUTACIONAL NO PROJETO AUTOMÁTICO DE REDES NEURAIS HÍBRIDAS E REDES NEUROFUZZY HETEROGÊNEAS**. 2000. Disponível em: <https://www.dca.fee.unicamp.br/~vonzuben/theses/emi_mest/cap2.pdf>. Acesso em: 19/08/2023.
- JAEGER DHAVAL TRIVEDI, M. S. H. **Mobile Device Voice Recordings at King's College London (MDVR-KCL) from both early and advanced Parkinson's disease patients and healthy controls**. 2019. Disponível em: <<https://zenodo.org/records/2867216>>. Acesso em: 19/08/2023.
- JANKOVIC, J. **Parkinson's disease: clinical features and diagnosis**. *Journal of neurology, neurosurgery, and psychiatry*, PubMed, v. 79(4), p. 368–376, 2008. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/18344392/>>. Acesso em: 07/04/2023.

JAYASWAL, V. **Performance Metrics: Confusion matrix, Precision, Recall, and F1 Score**. 2020. Disponível em: <<https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262>>. Acesso em: 21/08/2023.

JONES, M. T. **Um mergulho profundo nas redes neurais recorrentes**. 2017. Disponível em: <[https://imasters.com.br/data/um-mergulho-profundo-nas-redes-neurais-recorrentes#:~:text=Uma%20rede%20neural%20recorrente%20\(RNN,alimenta%C3%A7%C3%A3o%20apenas%20para%20camadas%20subsequentes\).](https://imasters.com.br/data/um-mergulho-profundo-nas-redes-neurais-recorrentes#:~:text=Uma%20rede%20neural%20recorrente%20(RNN,alimenta%C3%A7%C3%A3o%20apenas%20para%20camadas%20subsequentes).>)>. Acesso em: 19/08/2023.

LAMBA TARUN GULATI, A. J. P. R. R. **A Speech-Based Hybrid Decision Support System for Early Detection of Parkinson's Disease**. *Arab J Sci Eng*, v. 48, p. 2247–2260, 2023. Disponível em: <<https://link.springer.com/article/10.1007/s13369-022-07249-8>>. Acesso em: 18/07/2023.

LOBO, L. C. **Inteligência artificial e medicina**. *Revista Brasileira de Educação Médica, SciELO Brasil*, v. 41, p. 185–193, 2017. Disponível em: <<https://www.scielo.br/j/rbem/a/f3kqKJjVQJxB4985fDMVb8b/?format=html>>. Acesso em: 07/04/2023.

MAXIME. **What is a Transformer?** 2019. Disponível em: <<https://medium.com/inside-machine-learning/what-is-a-transformer-d07dd1fbec04>>. Acesso em: 19/08/2023.

ORACLE. **O que é Deep Learning?** 2023. Disponível em: <<https://www.oracle.com/br/artificial-intelligence/machine-learning/what-is-deep-learning/>>. Acesso em: 19/08/2023.

SAÚDE, M. D. **Doença de Parkinson**. 2019. Disponível em: <<https://bvsmis.saude.gov.br/doenca-de-parkinson/>>. Acesso em: 07/04/2023.

SAÚDE, M. D. **A avaliação do paciente em cuidados paliativos**. *Biblioteca Virtual em Saúde Prevenção e Controle de Câncer*, v. 1, p. 286, 2022. Disponível em: <https://www.inca.gov.br/sites/ufu.sti.inca.local/files/media/document/completo_serie_cuidados_paliativos_volume_1.pdf>. Acesso em: 18/07/2023.

SERVICES, A. W. **O que é regressão logística?** 2023. Disponível em: <<https://aws.amazon.com/pt/what-is/logistic-regression/#:~:text=A%20regress%C3%A3o%20log%C3%ADstica%20%C3%A9%20uma,resultados%2C%20como%20sim%20ou%20n%C3%A3o.>>>. Acesso em: 21/08/2023.

SERVICES, A. W. **O que é uma rede neural?** 2023. Disponível em: <<https://aws.amazon.com/pt/what-is/neural-network/#:~:text=Uma%20rede%20neural%20%C3%A9%20um,camadas%2C%20semelhante%20ao%20c%C3%A9rebro%20humano.>>>. Acesso em: 21/08/2023.

UKANI, V. **Parkinson's Disease Data Set**. 2020. Disponível em: <<https://www.kaggle.com/datasets/vikasukani/parkinsons-disease-data-set>>. Acesso em: 19/08/2023.

WANG YAN FU, B. S. L. C. K. R. Z. C.-Y. L. . Q. **Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset**. *Front Aging Neurosci*, v. 14, p. 1036588, 2022. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9691375/>>. Acesso em: 18/07/2023.

WANG YAN FU, B. S. L. C. K. R. Z. C.-Y. L. Q. **Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset**. *Front Aging Neurosci*, v. 14, 2022. Disponível em: <<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9691375/>>. Acesso em: 07/04/2023.

YOUSIF HOSSAM MAGDY BALAHA, A. Y. H. E. M. E.-G. N. R. **A generic optimization and learning framework for Parkinson disease via speech and handwritten records.** *J Ambient Intell Human Comput*, v. 14, p. 10673–10693, 2023. Disponível em: <<https://link.springer.com/article/10.1007/s12652-022-04342-6>>. Acesso em: 18/07/2023.

YüCELBAŞ, C. **A new approach: information gain algorithm-based k-nearest neighbors hybrid diagnostic system for Parkinson's disease.** *Physical and Engineering Sciences in Medicine*, v. 44, p. 511–524, 2021. Disponível em: <<https://doi.org/10.1007/s13246-021-01001-6>>. Acesso em: 18/07/2023.

ZAPAROLLI, D. **A inteligência artificial chega à saúde.** 2022. Disponível em: <<https://revistapesquisa.fapesp.br/a-inteligencia-artificial-chega-a-saude/>>. Acesso em: 19/08/2023.