

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
CAMPUS DE BAURU
Faculdade de Ciências
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

ANDRÉ LUIZ DA SILVA JUNIOR

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

Bauru
2023

André Luiz da Silva Junior

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

Trabalho de conclusão do curso do 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

Bauru
2023

S586a

Silva Junior, André Luiz da

ABORDAGEM DE APRENDIZADO PROFUNDO PARA
CLASSIFICAÇÃO DE PARKINSON POR MEIO DE SINAIS DE
VOZ / André Luiz da Silva Junior. -- Bauru, 2023

75 p.

Trabalho de conclusão de curso (Bacharelado - Ciência da
Computação) - Universidade Estadual Paulista (Unesp), Faculdade de
Ciências, Bauru

Orientador: Clayton Reginaldo Pereira

1. Doença de Parkinson. 2. Aprendizado de Máquina. 3.
Aprendizado Profundo. 4. Inteligência Artificial. 5. Redes Neurais. I.

Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de
Ciências, Bauru. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

ANDRÉ LUIZ DA SILVA JUNIOR

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

Trabalho de conclusão do curso do 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 .

Banca Examinadora

Profº Dr. Clayton Reginaldo Pereira
Orientador
Universidade Estadual Paulista "Júlio de Mesquita
Filho"
Faculdade de Ciências
Departamento de Computação

Profº Dra. Simone das Graças Domingues Prado
Universidade Estadual Paulista "Júlio de Mesquita
Filho"
Faculdade de Ciências
Departamento de Computação

Profº Dr. Douglas Rodrigues
Universidade Estadual Paulista "Júlio de Mesquita
Filho"
Faculdade de Ciências
Departamento de Computação

Bauru , 17 de Novembro de 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 aplicadas a três conjuntos de dados com características próprias para um mais amplo estudo dos casos, técnicas de aprendizado de máquina. Dois desses conjuntos de dados foram analisados utilizando os métodos clássicos como a regressão logística, máquina de vetores de suporte, k-ésimo vizinho mais próximo, floresta aleatória, árvore de decisão, *Naive Bayes* e *XGBoost*, e um com a utilização de aprendizado de máquina com séries temporais, utilizado técnicas de aprendizado profundo como a rede neural convolucional e a rede neural recorrente, após o treino o modelo com o melhor desempenho foi selecionado para cada dataset. Os resultados revelaram que o modelo que teve um desempenho melhor nos dois conjunto de dados que utilizaram aprendizado de máquina clássico foi a Regressão Logística, e no aprendizado profundo 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, Aprendizado de Máquina, Aprendizado Profundo, Rede Neural Convolucional, Rede Neural Recorrente.

ABSTRACT

Parkinson's disease is a progressive neurological disorder that affects the brain. The development of specialized systems capable of automatically and with high accuracy diagnosing early stages of Parkinson's disease based on speech signals would represent an important contribution to the healthcare sector. To this end, machine learning techniques were applied to three sets of data with specific characteristics for a broader study of the cases. Two of these data sets were analyzed using classical methods such as logistic regression, support vector machine, k-th nearest neighbor, random forest, decision tree, Naive Bayes and XGBoost, and one using machine learning with time series, using deep learning techniques such as the convolutional neural network and the recurrent neural network, after training the model with the best performance was selected for each dataset. The results revealed that the model that performed best on the two datasets that used classical machine learning was Logistic Regression, and in deep learning it was Convolutional Neural Network. The results obtained revealed a classification accuracy of 92.31% and 86.84% for the two sets of data, which highlights the ability to obtain excellent results even with a limited set of data. Furthermore, accuracies of 91.43% and 86.72%, recalls of 100.00% and 97.37%, and F1-Scores of 95.52% and 91.74% were observed. These values corroborate the quality of the results achieved. In turn, the Convolutional Neural Network presented a classification accuracy of 86.69%, precision of 86.83%, recall of 88.20% and F1-Score of 87.51%, even with a number reduced training times. This highlights the effectiveness of this model as a valuable alternative for classification. These results indicate that rankings will play a significant role in advancing the medical sector. It is important to highlight the importance of careful analysis when choosing the model to be used, as different sets of data may present more favorable results in different models.

KEY-WORDS: Parkinson's Disease, Diagnosis, Machining learning, Deep learning, Convolutional Neural Network, Recurrent Neural Network.

LISTA DE FIGURAS

1	Redes Neurais em Camadas	21
2	Redes Neurais Recorrentes	23
3	Exemplo de <i>Support Vector Machine</i>	26
4	Distribuição do Primeiro <i>Dataset</i>	38
5	Primeira Distribuição de <i>Features</i> do Primeiro <i>Dataset</i>	39
6	Segunda Distribuição de <i>Features</i> do Primeiro <i>Dataset</i>	39
7	Análise T-SNE do Primeiro <i>Dataset</i>	40
8	Distribuição do Segundo <i>Dataset</i>	40
9	Distribuição de Gênero Segundo <i>Dataset</i>	41
10	Distribuição do Número de Pulsações do Segundo <i>Dataset</i>	41
11	Distribuição do Número de Pulsações médias do Segundo <i>Dataset</i>	42
12	Distribuição do Terceiro <i>Dataset</i> considerando a duração	42
13	Distribuição do Terceiro <i>Dataset</i> sem considerar a duração	43
14	Gráfico de Correlação do Primeiro <i>Dataset</i>	44
15	<i>Features</i> com mais Correlações do Primeiro <i>Dataset</i>	45
16	Vâriância dos Componentes de PCA do Segundo <i>Dataset</i>	45
17	<i>Scatterplot</i> do PCA dos dois principais <i>features</i> selecionados do Segundo <i>Dataset</i>	47
18	Dado Temporais	47
19	Transformadas Fourier	48
20	Coeficientes <i>Filter Bank</i>	48
21	<i>Mel Frequency Cepstral Coefficients</i>	48
22	Tabela criada	49
23	Configuração do <i>XGBoost</i>	51
24	Regressão Linear no Primeiro <i>Dataset</i>	52
25	Regressão Linear no Segundo <i>Dataset</i>	53
26	Máquina de vetores de suporte no Primeiro <i>Dataset</i>	54
27	Máquina de vetores de suporte no Segundo <i>Dataset</i>	55
28	<i>K-Neighbors</i> no Primeiro <i>Dataset</i>	56
29	<i>K-NN</i> no Segundo <i>Dataset</i>	57
30	Floresta Aleatória no Primeiro <i>Dataset</i>	58
31	Floresta Aleatória no Segundo <i>Dataset</i>	59
32	<i>Decision Tree</i> no Primeiro <i>Dataset</i>	60
33	Árvore de Decisão no Segundo <i>Dataset</i>	61
34	<i>Naive Bayes</i> no Primeiro <i>Dataset</i>	62
35	<i>Naive Bayes</i> no Segundo <i>Dataset</i>	63
36	<i>XGBoost</i> no Primeiro <i>Dataset</i>	64
37	<i>XGBoost</i> no Segundo <i>Dataset</i>	65

38	Treinamento CNN	65
39	Acurácia do Treinamento CNN	66
40	Perda do Treinamento CNN	66
41	Matriz de Confusão CNN	66
42	Configuração da Rede de Convolução	67
43	Treinamento RNN	67
44	Acurácia do Treinamento RNN	67
45	Perda do Treinamento RNN	68
46	Matriz de Confusão RNN	68
47	Configuração da Rede Recorrente	69

LISTA DE QUADROS

Quadro 1 – Testes de Acurácia	46
Quadro 2 – Resultados do primeiro <i>dataset</i>	69
Quadro 3 – Resultados do segundo <i>dataset</i>	69
Quadro 4 – Resultados do terceiro <i>dataset</i>	70

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	13
1.1	PROBLEMA	13
1.2	JUSTIFICATIVA	14
1.3	OBJETIVOS	14
1.3.1	OBJETIVOS GERAIS	14
1.3.2	OBJETIVOS ESPECÍFICOS	14
1.4	METODOLOGIA DE PESQUISA	15
2	REVISÃO TEÓRICA	16
2.1	Doença de Parkinson	16
2.1.1	Introdução	16
2.1.2	Sintomas	16
2.1.3	Deteção	17
2.1.4	Coleta de Dados	17
2.1.5	Necessidade de métodos de IA para deteção de Parkinson	18
2.2	Aprendizado de Máquina	20
2.2.1	Aprendizado Profundo	20
2.2.2	Redes Neurais	20
2.2.2.1	Redes Neurais Multicamada	20
2.2.2.2	Redes Neurais Convolucionais	22
2.2.2.3	Redes Neurais Recorrentes	23
2.2.3	Atenção	25
2.2.4	Transformers	25
2.2.5	Máquina de Vetores de Suporte	26
2.2.6	Árvore de Decisão	27
2.2.7	Floresta Aleatória	28
2.2.8	Regressão Logística	28
2.2.9	XGBoost	29
2.2.10	Naive Bayes	30
2.2.11	K-ésimo Vizinho mais Próximo	30
3	TRABALHOS CORRELATOS	32
3.1	Inteligência Artificial e Medicina	32
3.2	A new approach: information gain algorithm-based k-nearest neighbors hybrid diagnostic system for Parkinson's disease	32
3.3	Early detection of Parkinson's disease from multiple signal speech: Based on Mandarin language dataset	33

3.4	A generic optimization and learning framework for Parkinson disease via speech and handwritten records	34
3.5	A Speech-Based Hybrid Decision Support System for Early Detection of Parkinson’s Disease	34
3.6	Early detection of Parkinson’s disease using machine learning	35
4	DATASETS	36
4.1	Parkinson’s Disease Data Set	36
4.2	Parkinson’s Disease (PD) classification	37
4.3	Mobile Device Voice Recordings at King’s College London (MDVR-KCL) from both early and advanced Parkinson’s disease patients and healthy controls . .	37
4.4	Análise dos dados	37
4.4.1	Primeiro Dataset	38
4.4.2	Segundo Dataset	40
4.4.3	Terceiro Dataset	42
4.5	Seleção de features	43
4.5.1	Primeiro Dataset	43
4.5.2	Segundo Dataset	45
4.5.3	Terceiro Dataset	47
5	TREINAMENTO	49
5.1	Métricas utilizadas	49
5.2	Sampling dos dados	50
5.3	Treinamento dos modelos	51
5.3.1	Regressão Logística	51
5.3.1.1	Primeiro Dataset	51
5.3.1.2	Segundo Dataset	52
5.3.2	Máquina de vetores de suporte	53
5.3.2.1	Primeiro Dataset	53
5.3.2.2	Segundo Dataset	54
5.3.3	K-NN	55
5.3.3.1	Primeiro Dataset	55
5.3.3.2	Segundo Dataset	56
5.3.4	Floresta Aleatória	57
5.3.4.1	Primeiro Dataset	57
5.3.4.2	Segundo Dataset	58
5.3.5	Árvore de decisão	59
5.3.5.1	Primeiro Dataset	59
5.3.5.2	Segundo Dataset	60
5.3.6	Naive Bayes	61
5.3.6.1	Primeiro Dataset	61

5.3.6.2	Segundo Dataset	62
5.3.7	XGBoost	63
5.3.7.1	Primeiro Dataset	63
5.3.7.2	Segundo Dataset	64
5.3.8	Deep Learning no Terceiro Dataset	65
5.3.8.1	CNN	65
5.3.8.2	RNN	67
5.4	Resultados	69
6	CONCLUSÃO	71
	REFERÊNCIAS	72

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 médica, 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 aprendizado profundo 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 aprendizado profundo 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 sútis 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 aprendizado profundo 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 aprendizado profundo 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 aprendizado de máquina 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 sintomas predominantes desta condição geralmente se manifestam por meio de tremores involuntários, sendo mais comuns nas mãos, braços, pernas e também na região da 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 IA (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 movimentos 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 IA 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 APRENDIZADO DE MÁQUINA

2.2.1 Aprendizado Profundo

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).

Por ser uma subárea de aprendizado de máquina pode ser utilizado tanto como aprendizados supervisionados quanto não supervisionados. O treinamento de redes neurais profundas também envolve a minimização de uma função de custo usando algoritmos de otimização como o gradiente descendente e suas variantes.

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.

Essa técnica 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.

2.2.2 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).

2.2.2.1 Redes Neurais Multicamada

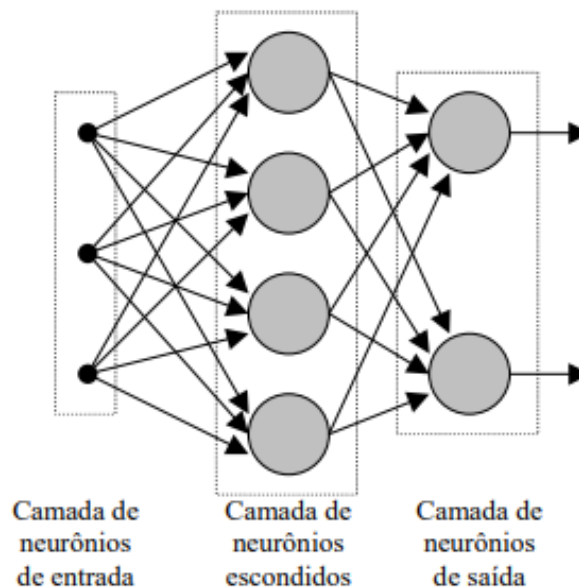
Redes neurais multicamadas são modelos avançados de aprendizado de máquina inspirados na estrutura do sistema nervoso humano, mais especificamente nos neurônios do cérebro. Essas redes consistem em várias camadas interconectadas de neurônios artificiais, permitindo um processamento mais complexo e uma 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).

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 carregada 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.

Figura 1 – Redes Neurais em Camadas



Fonte: IYODA(2000)

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 pelas equações 2.1 e 2.2:

$$z_j^{(l)} = \sum_{i=1}^{N^{l-1}} w_{ji}^{(l)} a_i^{(l-1)} + b_j^{(l)} \quad (2.1)$$

$$a_j^{(l)} = g(z_j^{(l)}) \quad (2.2)$$

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.

2.2.2.2 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 com a utilização da equação 2.3:

$$z(i, j) = (X * W)(i, j) = \sum_m \sum_n X(m, n) * W(i - m, j - n) \quad (2.3)$$

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 com a equação 2.4 em um janela de tamanho $f \times f$:

$$Y(i, j) = \max_{m,n} X(f * i + m, f * j + n) \quad (2.4)$$

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.

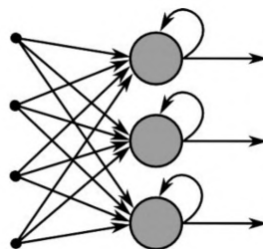
O treinamento das CNNs é realizado por meio da apresentação de um grande conjunto de dados de treinamento, em que os pesos das conexões são ajustados para minimizar o erro entre as saídas previstas e as saídas reais. A popularidade das CNNs cresceu significativamente devido à sua eficácia em muitas tarefas 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.

2.2.2.3 Redes Neurais Recorrentes

As Redes Neurais Recorrentes(RNN) 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



Fonte: ACADEMY(2022)

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 (ACADEMY, 2022a).

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):

$$h_t = f(W_{hx} + W_{hh} * h_{t-1} + b_h) \quad (2.5)$$

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 é um algoritmo utilizado no treinamento de redes neurais artificiais. Ele consiste em retroceder o erro calculado na saída da rede neural e ajustar os pesos das conexões entre os neurônios em camadas anteriores, permitindo assim que a rede aprenda a partir dos dados de treinamento. O processo envolve a propagação do erro da camada de saída para as camadas intermediárias, utilizando técnicas de otimização, como o gradiente descendente, para minimizar o erro global da rede (ACADEMY, 2022b). A retropropagação através do tempo é uma aplicação de retropropagação para sequenciar modelos com um estado oculto.

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 *Long Short Term Memory* (LSTM) e *Gated Recurrent Unit* (GRU), 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.

2.2.3 Atenção

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 RNN. 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 RNNs 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 cálculo de atenção é feito com a função *softmax*, que é descrita pela equação 2.6:

$$a_t(i) = \frac{\exp(e_t(i))}{\sum_{j=1}^n \exp(e_t(j))} \quad (2.6)$$

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:

$$c_t = \sum_{i=1}^n a_t(i) * h_i \quad (2.7)$$

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 autoatenção, que permite ao modelo ponderar a importância de diferentes partes da entrada durante a geração de uma saída, 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. 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).

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.

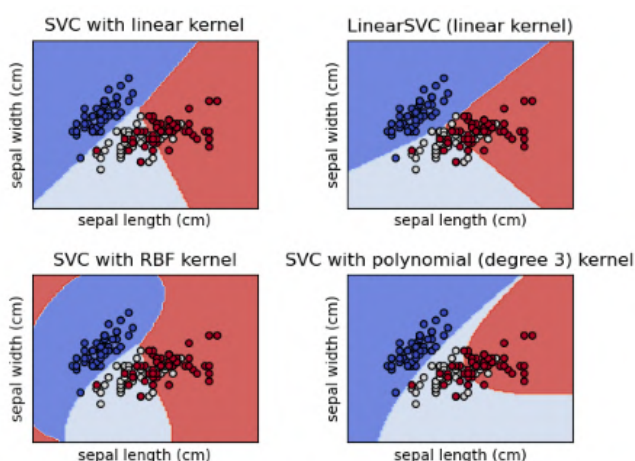
Os *transformers* demonstraram serem poderosos em diversas aplicações, desde tradução automática e sumarização de texto até resposta a perguntas e tarefas de visão computacional. Modelos pré-treinados baseados em *transformers*, como o BERT da *google* e o chat GPT da *open AI*, atingiram resultados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

2.2.5 Máquina de Vetores de Suporte

As máquinas de vetores de suporte (SVM) 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*



Fonte: DEVELOPERS(2023)

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 que classificam instâncias em duas classes distintas com o SVM encontrando um hiperplano de decisão que separa as duas classes de maneira ótima. Esse hiperplano é escolhido de forma que a margem (distância entre o hiperplano e os pontos mais próximos de cada classe) seja maximizada, SVMs de classificação multi-classe que estendem o SVM para lidar com mais de duas classes, podem ser um contra todos que treinam um classificador SVM para cada classe, distinguindo-a das demais, durante a previsão, escolhe a classe com a pontuação mais alta entre todos os classificadores ou um contra um onde cria um classificador SVM para cada par de classes, durante a previsão, a classe mais frequentemente prevista é escolhida ou ainda o hierárquico que constrói uma hierarquia de classificadores binários que são treinados para distinguir entre subconjuntos de classes, e SVMs de regressão, que predizem valores numéricos em vez de classes, em vez de encontrar um hiperplano que separe classes, o SVM de regressão busca encontrar um hiperplano que melhor se ajuste aos dados, minimizando as penalidades para as instâncias que ficam fora da margem de tolerância.

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. Nele o parâmetro de regularização C é introduzido para controlar a quantidade de penalização por erros, C é um hiperparâmetro que determina o equilíbrio entre maximizar a margem e minimizar a classificação incorreta. Um C menor permite uma margem mais larga, aceitando alguns erros de classificação (KARIMI, 2022).

2.2.6 Árvore de Decisão

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.

A árvore começa com o conjunto de dados completo. No processo de construção, ela escolhe a

melhor característica para dividir os dados em subconjuntos. Essa escolha é feita com base em critérios como *Gini impurity* para classificação ou redução de variância para regressão. A raiz representa a característica que melhor divide o conjunto de dados. Os nós subsequentes representam características que dividem os dados nos subconjuntos mais puros possível, ao percorrer a árvore, cada nó interno representa uma decisão com base no valor de uma característica específica. As arestas indicam os possíveis resultados dessa decisão.

As folhas representam os resultados finais ou as classes previstas para uma determinada instância de dados. Cada caminho da raiz até uma folha representa uma sequência de decisões que levam a esse resultado.

2.2.7 Floresta Aleatória

A Floresta Aleatória 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, a Floresta Aleatória 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.

A Floresta Aleatória é amplamente utilizado em diversas áreas, incluindo classificação de imagem, detecção de fraudes, bioinformática, processamento de linguagem natural, entre outras. Embora a Floresta Aleatória 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.

A aplicação utiliza a função logística ou função sigmoide à combinação linear das variáveis independentes. A função sigmoide tem a forma como se vê na equação 2.8, onde z é a combinação linear das variáveis.

$$f(x) = \frac{1}{1 + e^{-z}} \quad (2.8)$$

A combinação linear é calculada como a soma ponderada das variáveis independentes, cada uma multiplicada por seu peso associado com a função logística transformando a combinação linear em uma probabilidade entre 0 e 1.

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.

A técnica constrói árvores de decisão sequencialmente, cada nova árvore corrige os erros das árvores anteriores.

O treinamento do *XGBoost* ocorre em etapas:

- Na inicialização se começa com uma predição simples como a média para problemas de regressão ou a proporção de classes para problemas de classificação.
- Calcula-se o gradiente e o Hessiano da função de objetivo em relação às predições atuais.
- Constrói-se uma nova árvore para corrigir os erros residuais.
- Aplica-se regularização à árvore.
- Atualiza as predições com as previsões da nova árvore, ajustando-as proporcionalmente para evitar superajuste

Por fim, a previsão do XGBoost é feita somando as previsões de todas as árvores ponderadas pelos seus coeficientes.

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). A fórmula geral do Teorema de Bayes é ilustrada pela equação 2.9.

$$P(C | X) = \frac{P(X | C) * P(C)}{P(X)} \quad (2.9)$$

Onde:

- $P(C | X)$ é a probabilidade condicional da classe C dada as características X.
- $P(X | C)$ é a probabilidade condicional das características X dada a classe C.
- $P(C)$ é a probabilidade da classe C.
- $P(X)$ é a probabilidade das características X, que pode ser normalizada como uma constante para tornar a fórmula mais simples.

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-ésimo Vizinho mais Próximo

O algoritmo K-ésimo Vizinho mais Próximo (k-NN) 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.

Na classificação, durante a fase de treinamento, o modelo armazena os exemplos de treinamento no espaço de características, quando uma nova instância que precisa ser classificada é apresentada, o algoritmo calcula a distância dessa instância em relação a todas as outras instâncias no conjunto de treinamento. A distância pode ser calculada usando diversas métricas, como distância Euclidiana ou distância de Manhattan. O algoritmo identifica os k vizinhos mais próximos da nova instância, onde k é um parâmetro definido previamente, os rótulos das instâncias vizinhas são considerados, e a nova instância é classificada com base em uma votação majoritária. A classe mais comum entre os k vizinhos é atribuída à nova instância.

Já na regressão, similar à classificação, o treinamento envolve armazenar os exemplos de treinamento no espaço de características, quando uma nova instância que precisa ter seu valor previsto é apresentada, o algoritmo calcula a distância dessa instância em relação a todas as outras instâncias no conjunto de treinamento.

Em vez de votação, para tarefas de regressão, o algoritmo calcula uma média ponderada dos valores alvo das instâncias vizinhas, onde o peso é determinado pela proximidade. Instâncias mais próximas têm mais peso na média.

3 TRABALHOS CORRELATOS

3.1 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.

3.2 A NEW APPROACH: INFORMATION GAIN ALGORITHM-BASED K-NEAREST NEIGHBORS HYBRID DIAGNOSTIC SYSTEM FOR PARKINSON'S DISEASE

O artigo discorre 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.

3.3 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.

3.4 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.

3.5 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.

3.6 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.

4.1 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.

4.2 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.

4.3 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)*.

4.4 ANALISE DOS DADOS

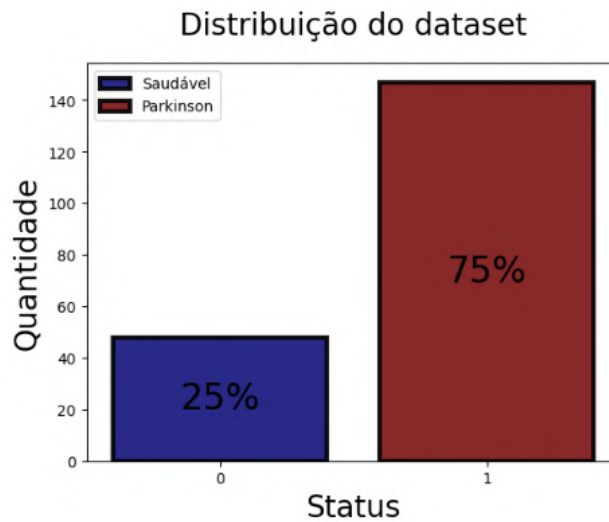
A fim de obter uma compreensão mais abrangente sobre a disposição dos dados nos conjuntos de dados, inicialmente foram realizadas análises abordando uma série de características. Essas análises

visaram avaliar a distribuição dos dados, verificar se ela está equitativa e investigar as relações entre as características e o diagnóstico de Parkinson.

4.4.1 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*.

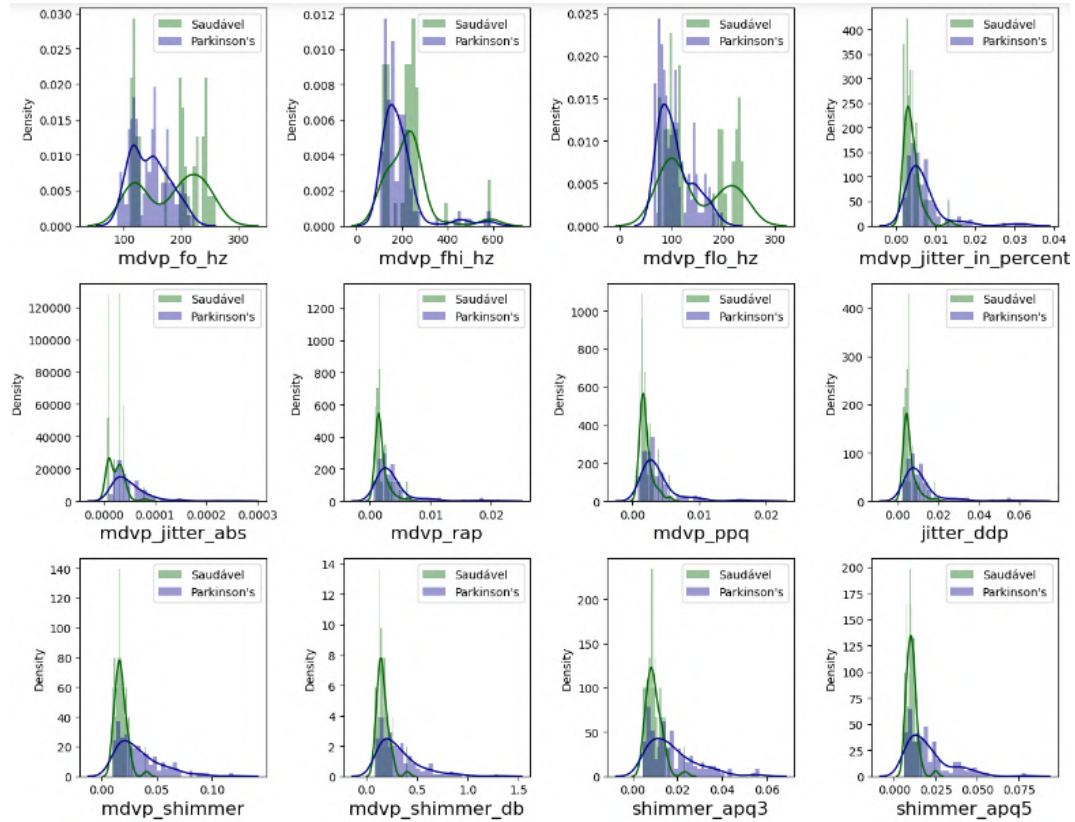
Figura 4 – Distribuição do Primeiro *Dataset*



Fonte: Elaborada pelo autor

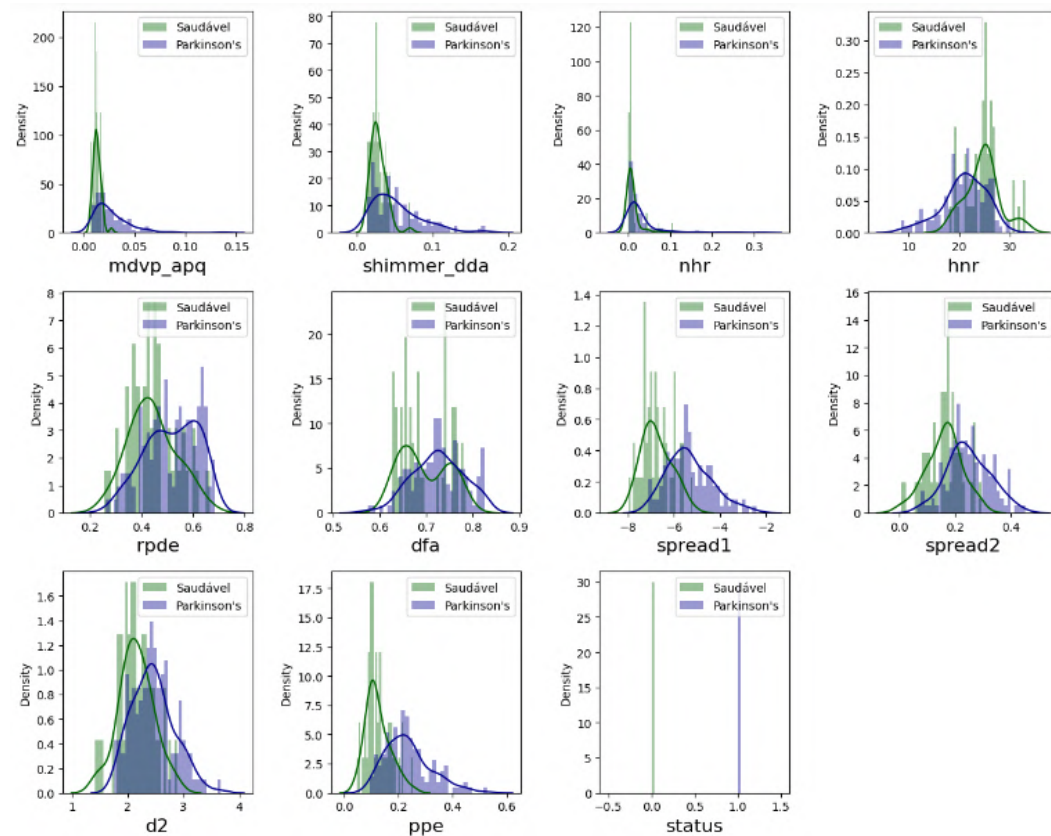
Posteriormente, examinamos a relação de cada característica quando aplicada em indivíduos saudáveis e naqueles com a condição da doença, notando uma distinta discrepância nos picos em cada gráfico, tornando a diferença entre eles evidente.

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



Fonte: Elaborada pelo autor

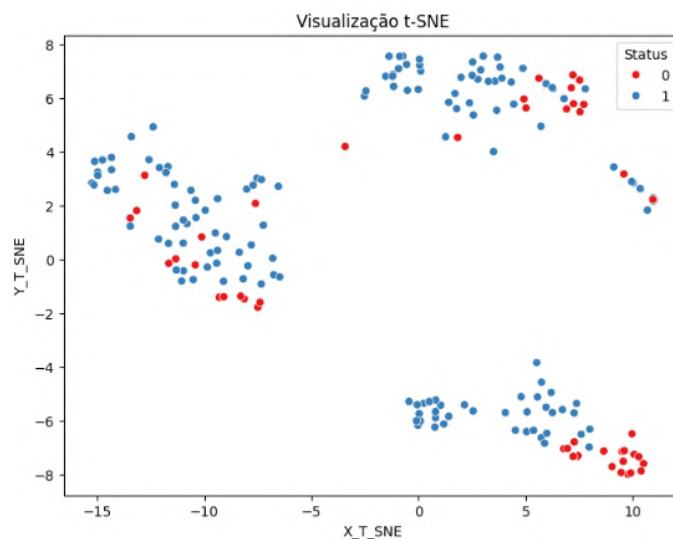
Figura 6 – Segunda Distribuição de *Features* do Primeiro *Dataset*



Fonte: Elaborada pelo autor

Ao analisar o gráfico do *scatterplot* de t-SNE, percebe-se 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*

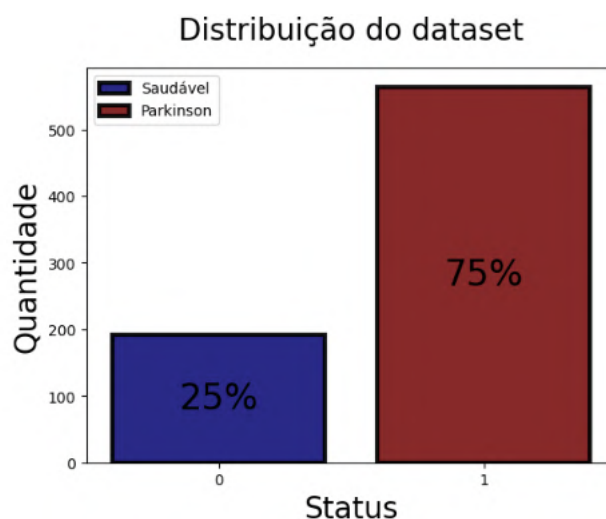


Fonte: Elaborada pelo autor

4.4.2 Segundo Dataset

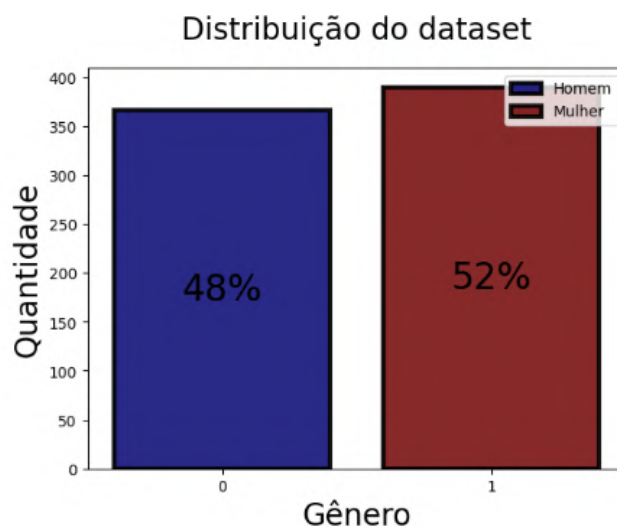
Da mesma forma que no primeiro conjunto de dados, a proporção entre gravações de áudio de indivíduos doentes e saudáveis permaneceu em 75% para 25%, mesmo diante da significativa ampliação da quantidade de dados disponíveis.

Figura 8 – Distribuição do Segundo *Dataset*



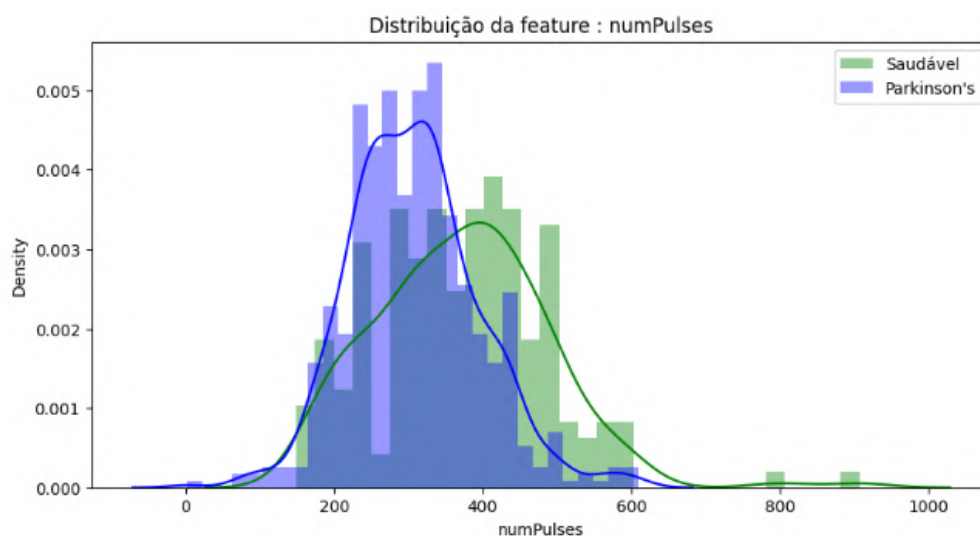
Fonte: Elaborada pelo autor

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*

Fonte: Elaborada pelo autor

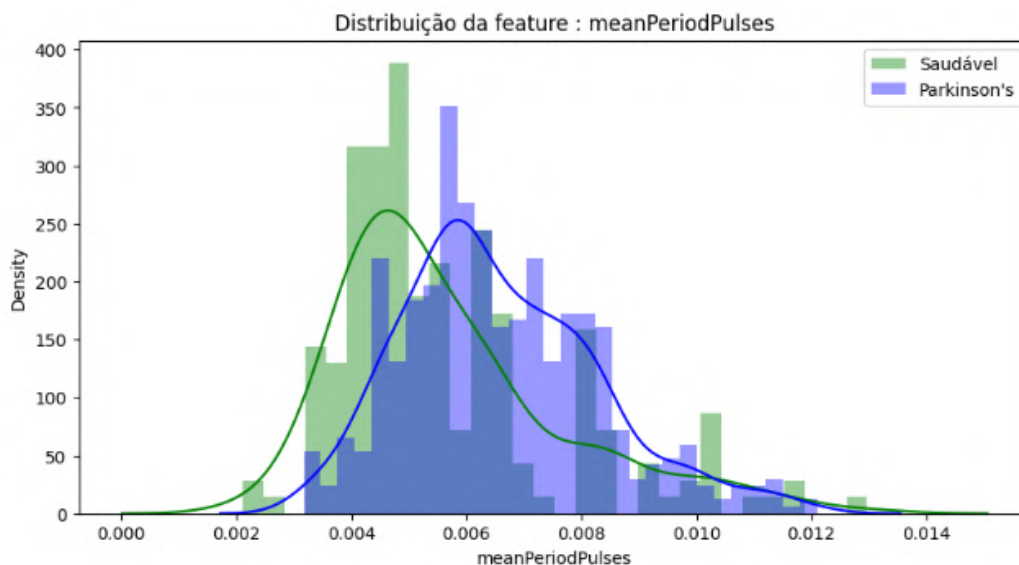
Nas distribuições dá para ver diferenças bem marcantes de quem tem a doença pra quem não tem, como por exemplo no numero de pulsações da onda sonora da voz sofre como pôde-se ver abaixo, vê-se que pessoas saudáveis tem um maior numero de pulsos na média.

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

Fonte: Elaborada pelo autor

Outra distribuição onde pôde-se ver a diferença é no periodo medio dos pulsos onde realmente é confirmado que a frequência dos pulsos é maior em pessoas saudáveis.

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

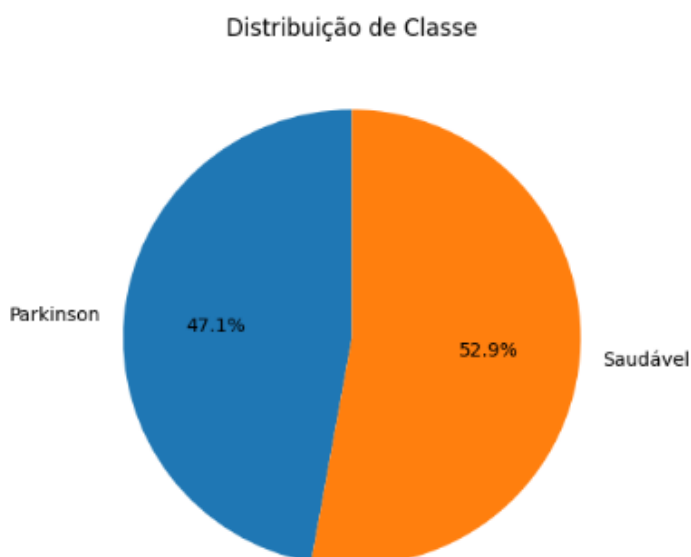


Fonte: Elaborada pelo autor

4.4.3 Terceiro Dataset

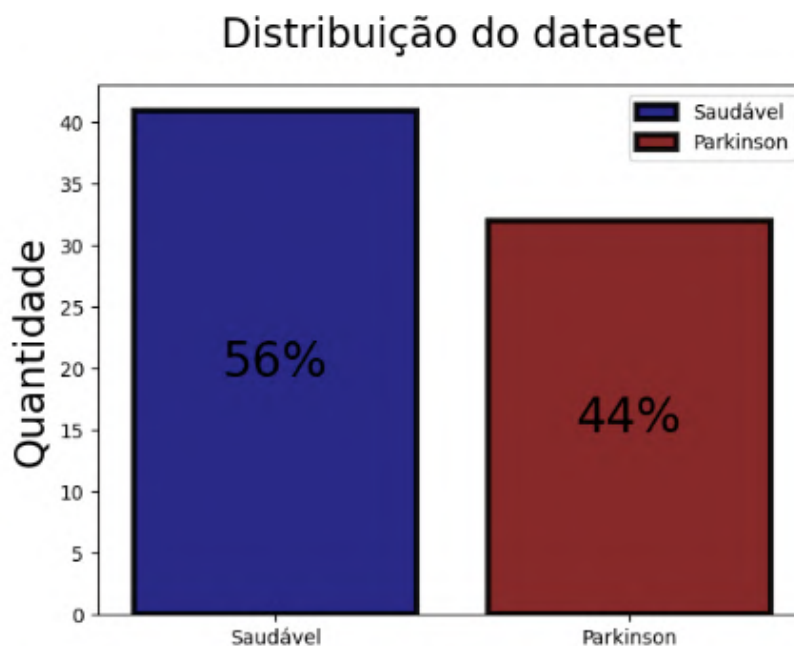
No conjunto de dados em questão, observa-se uma distribuição mais equilibrada entre indivíduos saudáveis e aqueles com Parkinson. Representando 47,1% dos áudios considerando sua duração, as gravações de pessoas afetadas pela condição são praticamente equivalentes às de indivíduos sem a doença, totalizando 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* considerando a duração



Fonte: Elaborada pelo autor

Já sem levar em consideração a duração dos áudios a distribuição do dataset continua bem equilibrado.

Figura 13 – Distribuição do Terceiro *Dataset* sem considerar a duração

Fonte: Elaborada pelo autor

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.

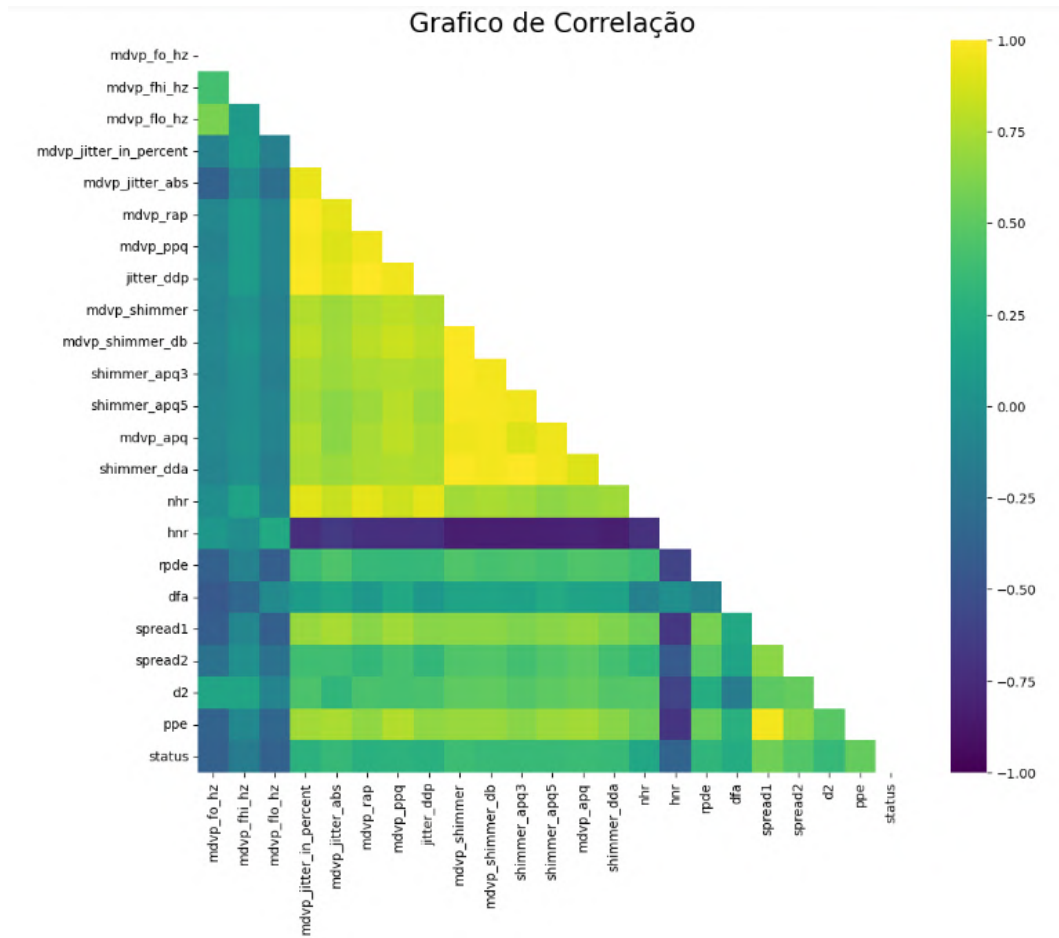
4.5 SELEÇÃO DE FEATURES

4.5.1 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.

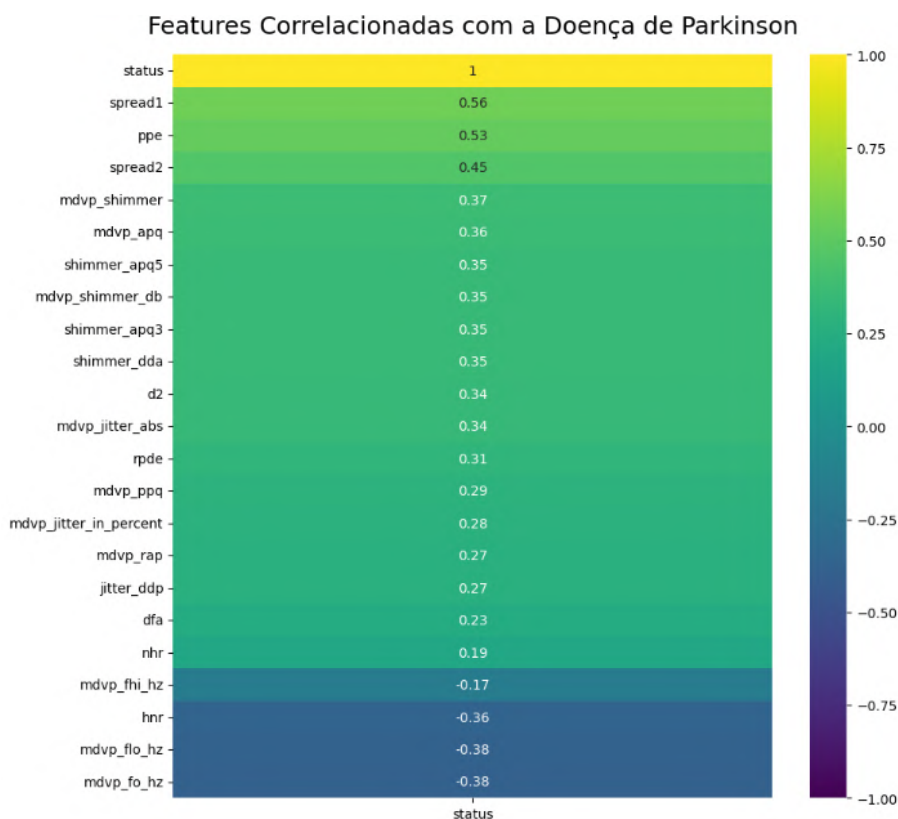
No gráfico da figura Figura 14, é possível observar as correlações mais notáveis, evidenciadas pelos tons mais intensos de amarelo e verde. Isso indica que MDVP_Shimmer, MDVP_Shimmer_dB, Shimmer_APQ3, Shimmer_APQ5, MDVP_APQ, e Shimmer_DDA exibem as correlações mais marcantes entre si. Além disso, MDVP_Jitter_in_Percent, MDVP_Jitter_Abs, MDVP_RAP, MDVP_PPQ, e Jitter_DDP também demonstram relações fortes.

Figura 14 – Gráfico de Correlação do Primeiro Dataset



Fonte: Elaborada pelo autor

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:

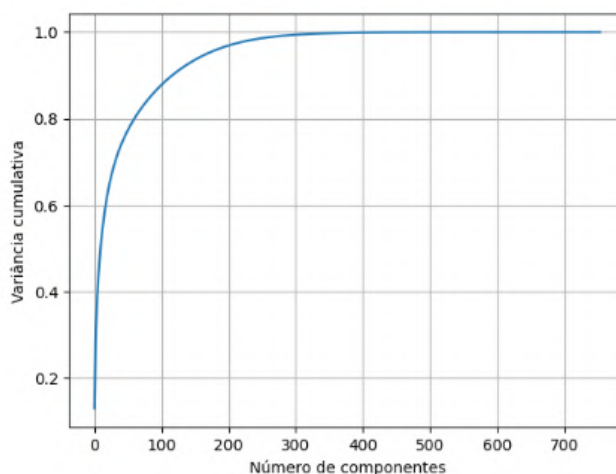
Figura 15 – *Features* com mais Correlações do Primeiro *Dataset*

Fonte: Elaborada pelo autor

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.

4.5.2 Segundo Dataset

Por esse *dataset* apresentar muitas *features* um método mais apropriado para reduzir e a dimensionalidade foi a utilização de PCA. Para saber quantas *features* selecionar foi analisado o gráfico da variância na figura Figura 16 para saber quando estava alta o bastante.

Figura 16 – Variância dos Componentes de PCA do Segundo *Dataset*

Fonte: Elaborada pelo autor

Pelo gráfico notamos que uma taxa superior a 95% só é alcançada em torno dos 200 *features*, porém ao realizar um teste de acurácia com poucas quantidades de *features* é verificado que consegue-se um bom valor com esses valores.

Quadro 1 – Testes de Acurácia

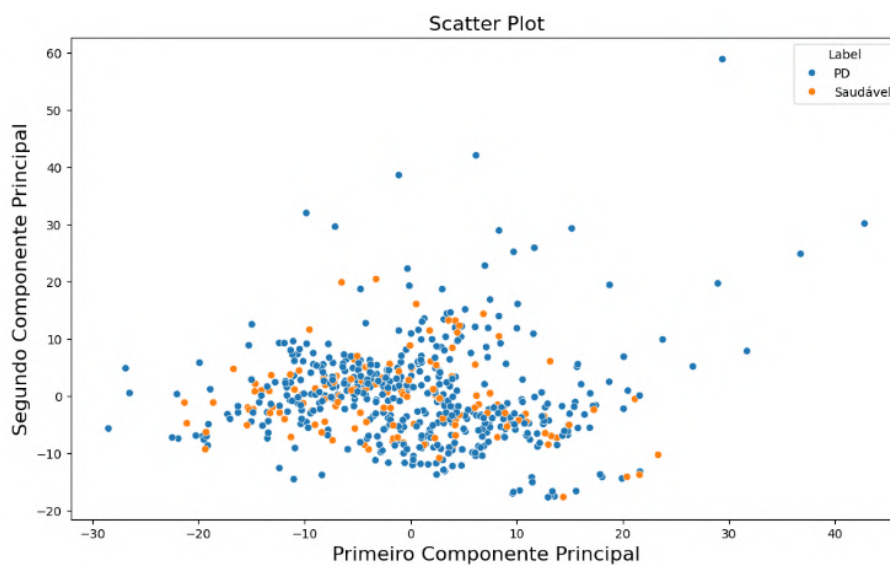
Componentes	Acurácia
1	78.94736642105263
2	83.55263157894737
3	79.605263157869474
4	83.55263157894737
5	84.210526315786947
6	86.8421052631579
7	83.55263157894737
8	83.55263157894737
9	83.55263157894737
10	83.55263157894737
11	82.89473684210526
12	80.92105263157895
13	83.55263157894737
14	83.55263157894737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

Fonte: Elaborada pelo autor

Pôde-se ver que com 6 *features* selecionadas temos o maior desempenho em acurácia, portanto será a utilizada para reduzir a dimensionalidade do *dataset*.

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

Figura 17 – *Scatterplot* do PCA dos dois principais *features* selecionados do Segundo *Dataset*



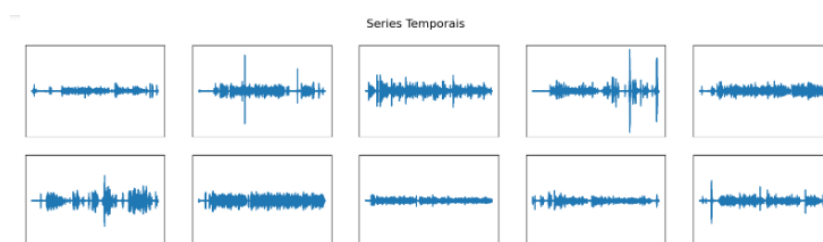
Fonte: Elaborada pelo autor

4.5.3 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.

Pôde-se visualizar as *features* extraídas dos áudio nos gráficos a seguir, na figura Figura 18 são demonstradas as séries temporais medidas ao longo do tempo, já na figura Figura 19 mostram as transformadas fourier que são a decomposição de um sinal em suas componentes elementares seno e cosseno, a figura Figura 20 demonstram os coeficientes filter bank que é uma matriz de filtros passa-faixa que separa o sinal de entrada em vários componentes e a figura Figura 21 são as mel frequency cepstral coefficients que são uma representação do espectro de potência de curto prazo de um som.

Figura 18 – Dado Temporais

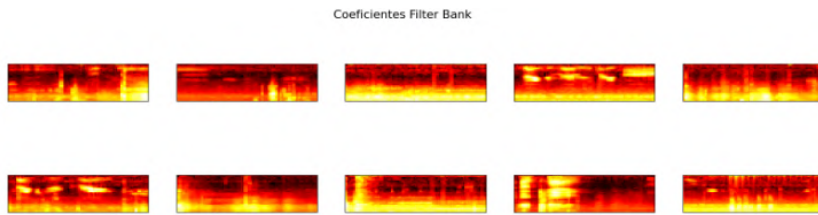


Fonte: Elaborada pelo autor

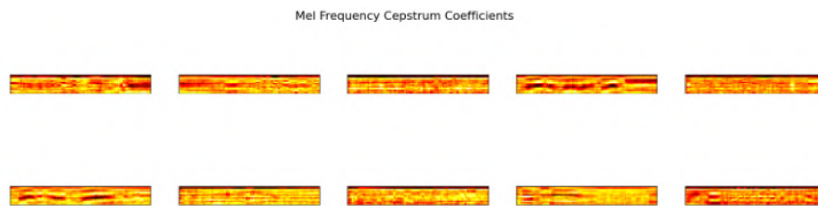
Figura 19 – Transformadas Fourier



Fonte: Elaborada pelo autor

Figura 20 – Coeficientes *Filter Bank*

Fonte: Elaborada pelo autor

Figura 21 – *Mel Frequency Cepstral Coefficients*

Fonte: Elaborada pelo autor

5 TREINAMENTO

Os dados que foram retirados da sessão anterior serão utilizados agora para o treinamento, além disso 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.

Figura 22 – Tabela criada

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

Fonte: Elaborada pelo autor

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 Matriz de Confusão, Acurácia, Precisão, *Recall* 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)

- **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.
- **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.
- **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:** A *Recall* 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 e (*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 $2 * (Precisão * Recall) / (Precisão + Recall)$. 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.

Serão explorados 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.

Na configuração é válido apontar que foram usadas as definições padrões de treino, com a *cross validation* com 10 *folds*, o que vale a pena destacar é a configuração da *XGBoost* que é apresentada na figura Figura 23.

Figura 23 – Configuração do *XGBoost*

```
import xgboost as xgb

clf = xgb.XGBClassifier(max_depth=3, n_estimators=300, learning_rate=0.05)
```

Fonte: Elaborada pelo autor

5.3.1 Regressão Logística

5.3.1.1 Primeiro Dataset

Aplicando a regressão logística no primeiro *dataset* foi verificado 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 precisão 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.

O valor de acurácia de treino 83,97% indica que o modelo conseguiu acertar a classe correta para cerca de 83,97% das instâncias durante o treinamento.

O *score* de *cross validation* fornece uma avaliação do desempenho do modelo em diferentes subconjuntos do conjunto de dados. Um valor de 82,63% sugere uma consistência razoável no desempenho do modelo em diferentes divisões do conjunto de dados.

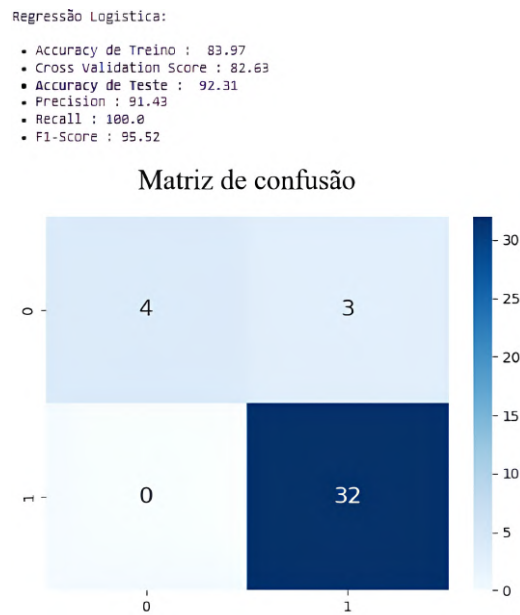
A acurácia de teste representa a capacidade do modelo de generalizar para dados não vistos. Uma acurácia de 92,31% indica que o modelo teve um bom desempenho na previsão de novos dados.

A precisão representa a proporção de instâncias positivas identificadas corretamente entre todas as instâncias previstas como positivas. A *recall* ou sensibilidade indica a proporção de instâncias positivas identificadas corretamente entre todas as instâncias positivas reais. Valores de 91,43% para precision e 100% para *recall* indicam uma capacidade sólida do modelo em identificar instâncias positivas, com poucos falsos positivos e nenhum falso negativo.

O F1-score é a média harmônica entre precisão e *recall*. Um valor de 95,52% sugere um equilíbrio entre *precision* e *recall*, indicando um bom desempenho geral do modelo na classificação das instâncias positivas.

A declaração de ausência de falsos negativos é uma observação importante, especialmente em contextos médicos, onde a identificação correta de casos positivos é crucial. A *recall* de 100% corrobora essa afirmação, indicando que o modelo conseguiu identificar todas as instâncias positivas presentes no conjunto de teste

Figura 24 – Regressão Linear no Primeiro *Dataset*



Fonte: Elaborada pelo autor

5.3.1.2 Segundo Dataset

Já no segundo *dataset*, foi obtido uma acurácia de treino de 81,62%, um *score* de *cross validation* de 80,78%, uma acurácia de teste de 86,84%, uma precisão de 86,72%, uma *recall* de 97,37% e um F1-*score* de 91,74%.

A acurácia de Treino de 81,62 indica eficiência em aprender padrões nos dados de treino. Já o *Score* de *cross validation* de 80,78% sugere generalização satisfatória para dados não vistos durante o treinamento.

A acurácia de teste de 86,84% destaca a capacidade de previsões precisas em um contexto real. A precisão de 86,72% sugere que a maioria das instâncias previstas como positivas realmente pertence à classe positiva. *Recall* de 97,37% contém a identificação correta da maioria das instâncias positivas, minimizando falsos negativos. O F1-*score* de 91,74% revela um equilíbrio sólido entre precisão e *recall*, consolidando o desempenho eficaz na detecção da condição alvo.

Figura 25 – Regressão Linear no Segundo *Dataset*

5.3.2 Máquina de vetores de suporte

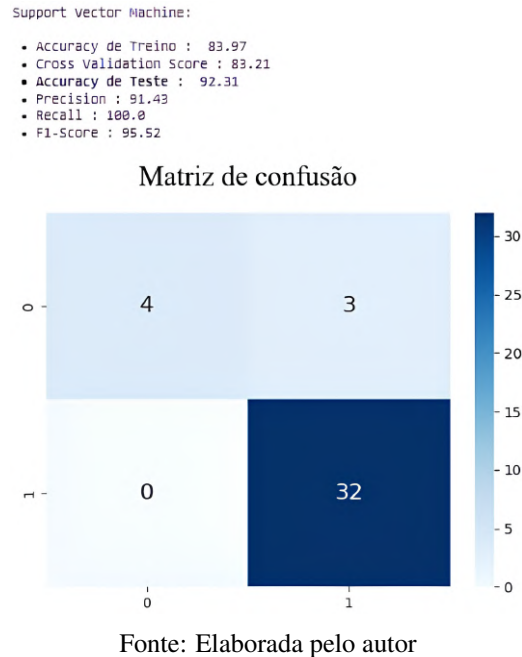
5.3.2.1 Primeiro Dataset

Agora para o Máquina de vetores de suporte no primeiro *dataset* conseguiu-se uma acurácia de treino de 83,97%, um *score* de *cross validation* de 83,21%, uma acurácia de teste de 92,31%, uma precisão de 91,43%, uma *recall* de 100% e um F1-*score* de 95,52%.

Com essa acurácia de treino, o modelo demonstra a capacidade de aprender eficientemente os padrões presentes nos dados de treinamento.

O *score* de *cross validation* próximo a 83,21% sugere que o SVM generaliza bem para dados não vistos durante o treinamento, indicando uma robustez na capacidade de previsão. A acurácia de teste de 92,31% destaca a habilidade do modelo em realizar previsões precisas em um cenário mais realista.

A precisão de 91,43% aponta que, das instâncias previstas como positivas, a grande maioria realmente pertencentes à classe positiva. A *recall* de 100% destaca que o SVM identificou corretamente todas as instâncias positivas, eliminando falsos negativos. O F1-*score* de 95,52% demonstra um equilíbrio entre precisão e *recall*, consolidando o bom desempenho do SVM na detecção da condição alvo. Esses resultados são semelhantes aos obtidos pela regressão logística, indicando que ambos os modelos são eficazes para este conjunto de dados específico.

Figura 26 – Máquina de vetores de suporte no Primeiro *Dataset*

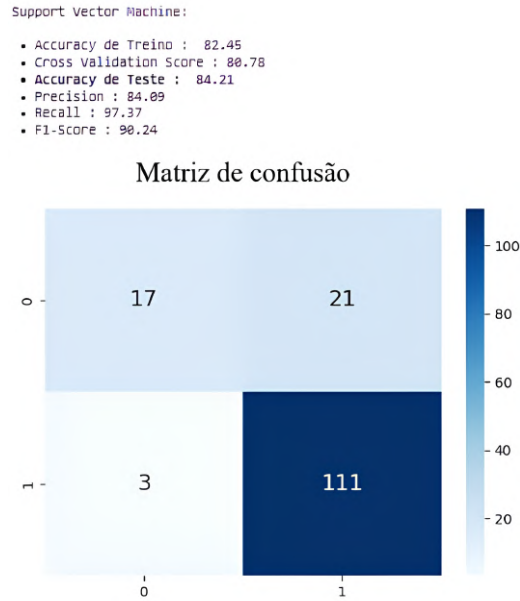
5.3.2.2 Segundo Dataset

E no segundo conseguiu-se uma acurácia de treino de 82.45%, um *score* de *cross validation* de 80.78%, uma acurácia de teste de 84.21%, uma precisão de 84.09%, uma *recall* de 97.37% e um *F1-score* de 90.24%.

Com uma acurácia de treino dessas demonstra-se a eficiência ao aprender padrões nos dados de treino. O *score* de *cross validation* sugere boa generalização para dados não vistos durante o treinamento. A acurácia de teste indica capacidade de realizar previsões precisas em um contexto mais realista.

A precisão revela que a maioria das instâncias previstas como positivas pertencentes à classe positiva e *arecall* de 97,37% demonstra identificação correta da grande maioria das instâncias positivas, minimizando falsos negativos. O *F1-score* mostra equilíbrio sólido entre precisão e *recall*, consolidando um desempenho eficaz.

Esses resultados indicam que o modelo SVM, embora tenha uma acurácia ligeiramente inferior em comparação com a regressão logística, ainda mantém um desempenho sólido na detecção da condição alvo no segundo conjunto de dados. A *recall* alta sugere a capacidade do SVM em identificar corretamente a grande maioria das instâncias positivas.

Figura 27 – Máquina de vetores de suporte no Segundo *Dataset*

Fonte: Elaborada pelo autor

5.3.3 K-NN

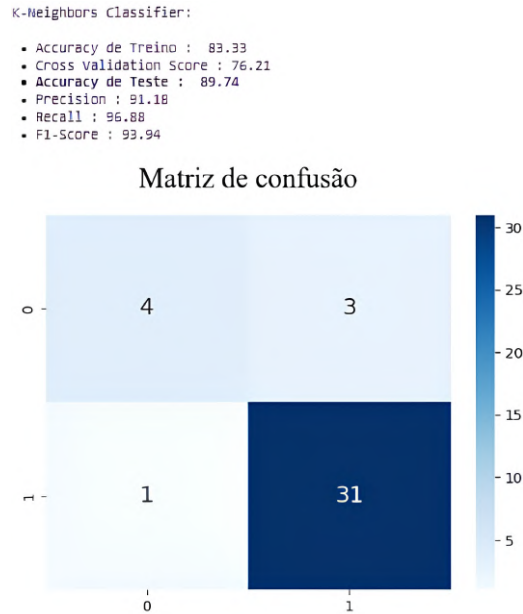
5.3.3.1 Primeiro Dataset

Com o K-NN no primeiro *dataset* teve-se uma acurácia de treino de 83,33%, um *score* de *cross validation* de 76.21%, uma acurácia de teste de 89,74%, uma precisão de 91.18%, uma *recall* de 96,88% e um F1-*score* de 93,94%.

A acurácia de Treino demonstrou-se eficaz ao aprender padrões nos dados de treino. O *score* de *cross validation* sugere uma generalização moderada para dados não vistos durante o treinamento, a acurácia de teste indica capacidade de realizar previsões precisas em um contexto mais realista. A precisão tem a maioria das instâncias previstas como positivas pertencentes à classe positiva. A *recall* identifica corretamente a grande maioria das instâncias positivas, minimizando falsos negativos. E o F1-*score* demonstra equilíbrio sólido entre precisão e *recall*, consolidando um desempenho eficaz.

Apesar de uma leve redução na acurácia de treino em comparação com os modelos anteriores, o K-Nearest Neighbors ainda mantém um desempenho robusto na detecção da condição alvo no primeiro conjunto de dados. Destaca-se a alta *recall*, indicando uma capacidade eficaz do modelo em identificar corretamente a grande maioria das instâncias positivas.

Figura 28 – K-Neighbors no Primeiro Dataset



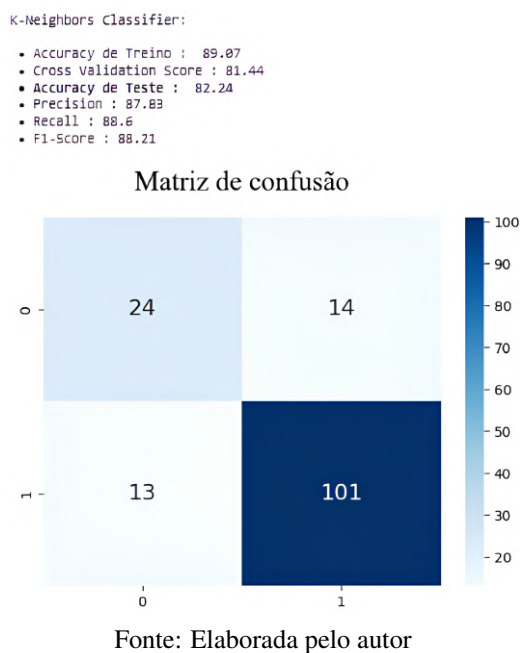
Fonte: Elaborada pelo autor

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 precisão de 87.83%, uma *recall* de 88.60% e um F1-score de 88.21%, observou-se mais uma diminuição da acurácia.

A acurácia de treino mostra eficiência em aprender padrões nos dados de treino. O *score* de *cross validation* sinaliza uma generalização razoável para dados não vistos durante o treinamento. A acurácia de teste indica capacidade de realizar previsões precisas no contexto realista do teste. A precisão tem a maioria das instâncias previstas como positivas pertencentes à classe positiva. A *recall* identifica corretamente a grande maioria das instâncias positivas, minimizando falsos negativos e o F1-score mantém um equilíbrio sólido entre precisão e *recall*, refletindo desempenho eficaz.

Apesar de uma redução na acurácia de treino em comparação com o primeiro conjunto de dados, o K-NN ainda demonstra um desempenho satisfatório na detecção da condição alvo no segundo conjunto.

Figura 29 – K-NN no Segundo *Dataset*

5.3.4 Floresta Aleatória

5.3.4.1 Primeiro Dataset

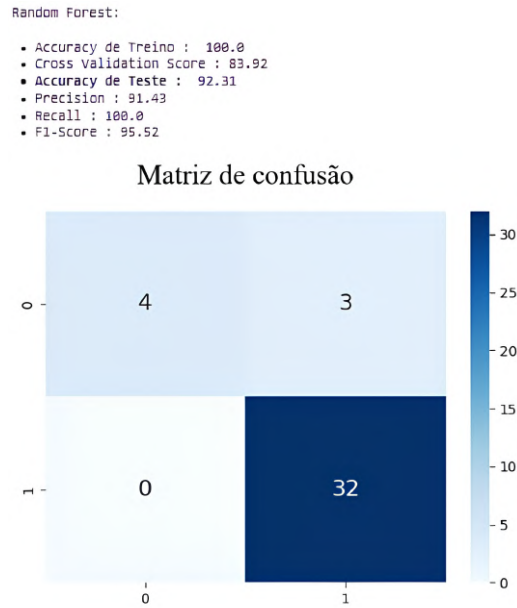
Utilizando Floresta Aleatória no primeiro *dataset* obteve-se uma acurácia de treino de 100,00%, um *score* de *cross validation* de 83,92%, uma acurácia de teste de 92,31%, uma precisão de 91.43%, uma *recall* de 100% e um F1-*score* de 95,52%.

A acurácia de treino destaca a capacidade do modelo de ajustar-se perfeitamente aos dados de treino. O *score* de *cross validation* tem boa generalização para dados não vistos durante o treinamento. A acurácia de teste tem alta precisão nas previsões em um contexto mais realista.

A precisão com a maioria das previsões positivas corretas. O *recall* identifica perfeitamente todas as instâncias positivas, minimizando os falsos negativos. E o F1-*score* tem um equilíbrio sólido entre precisão e *recall*, indicando um desempenho global eficaz.

Os resultados sugerem que o modelo Random Forest atingiu uma precisão notável na detecção da condição alvo no primeiro conjunto de dados, mantendo taxas extremamente altas de acertos tanto no treino quanto no teste.

Figura 30 – Floresta Aleatória no Primeiro Dataset



Fonte: Elaborada pelo autor

5.3.4.2 Segundo Dataset

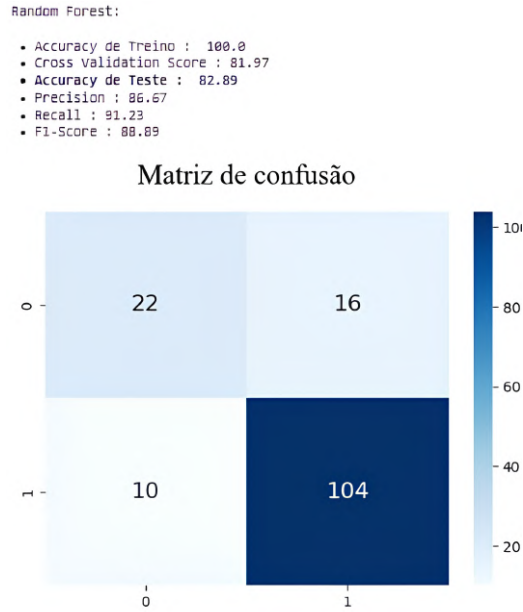
No segundo *dataset* conseguiu-se a acurácia de treino de 100.00%, um *score* de *cross validation* de 81.97%, uma acurácia de teste de 82.89%, uma precisão de 86.67%, uma *recall* de 91.23% e um *F1-score* de 88.89%, mantendo menores taxas que a regressão logística.

A acurácia de treino reflete uma adaptação perfeita do modelo aos dados de treino. O *score* de *cross validation* contém boa generalização para dados não utilizados no treinamento. A acurácia de teste tem alta precisão nas previsões em um contexto mais realista.

A precisão tem a maioria das previsões positivas corretas. A *recall* tem identificação eficaz da maioria das instâncias positivas, minimizando os falsos negativos. E o *F1-score* também tem um equilíbrio satisfatório entre precisão e *recall*.

Os resultados sugerem embora tenha atingido altos níveis de acurácia de treino, manteve um desempenho um pouco inferior ao primeiro conjunto de dados e comparável ao SVM e K-NN no segundo conjunto de dados.

Figura 31 – Floresta Aleatória no Segundo Dataset



Fonte: Elaborada pelo autor

5.3.5 Árvore de decisão

5.3.5.1 Primeiro Dataset

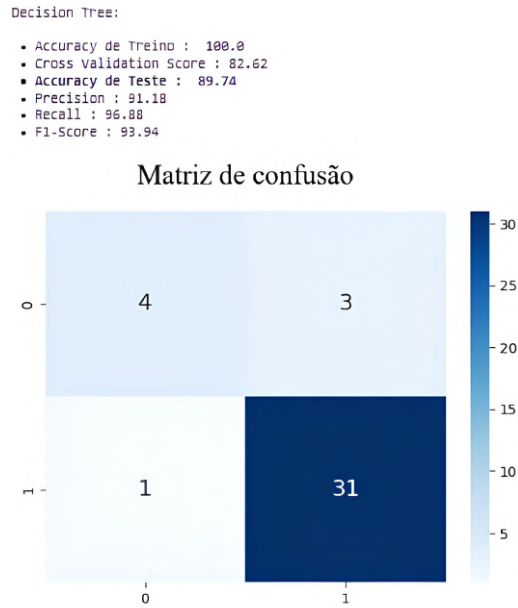
A Árvore de decisão para o primeiro alcançou uma acurácia de treino de 100,00%, um *score* de *cross validation* de 82,62%, uma acurácia de teste de 89,74%, uma precisão de 91,18%, uma *recall* de 96,88% e um *F1-score* de 93,94%.

A acurácia de Treino revela adaptação perfeita aos dados utilizados no treinamento. O *score* de *cross validation* tem boa capacidade de generalização para dados não vistos durante o treinamento. A acurácia de teste tem precisão elevada nas previsões em um contexto realista.

A precisão tem a maioria das previsões positivas corretas. A *Recall* alta identifica eficácia da maioria das instâncias positivas. E o *F1-score* tem o equilíbrio sólido entre precisão e *recall*.

Os resultados sugerem que a árvore de decisão apresenta um desempenho comparável ao *K-Neighbors* no primeiro conjunto de dados. Ambos os modelos demonstraram capacidade de adaptação aos dados de treino, boa generalização e altas taxas de acerto nas previsões de teste.

Figura 32 – *Decision Tree* no Primeiro Dataset



Fonte: Elaborada pelo autor

5.3.5.2 Segundo Dataset

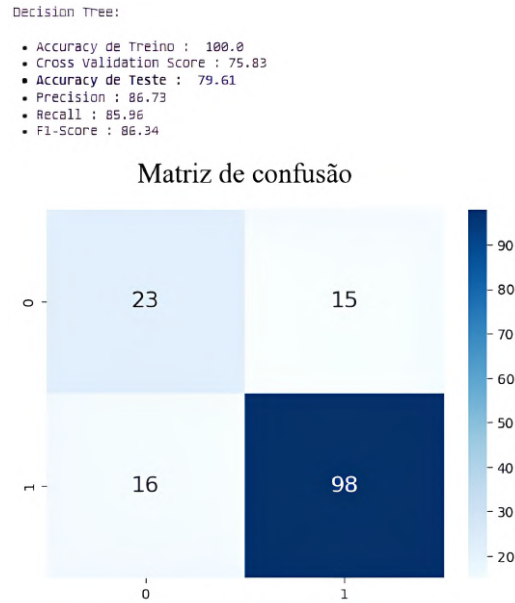
O segundo alcançou a acurácia de treino de 100.00%, um *score* de *cross validation* de 75.83%, uma acurácia de teste de 79.61%, uma precisão de 86.73%, uma *recall* de 85.96% e um F1-score de 86.34%.

A acurácia de treino indica adaptação perfeita aos dados de treino. O *score* de *cross validation* tem uma menor capacidade de generalização para dados não vistos durante o treinamento. A acurácia de teste tem precisão reduzida nas previsões em um contexto mais realista.

A precisão tem a maioria das previsões positivas corretas. A *Recall* identifica com eficácia a maioria das instâncias positivas, minimizando falsos negativos. E o F1-score apresenta boa harmonia entre precisão e *recall*, indicando desempenho razoável.

Os resultados revelam que a árvore de decisão no segundo conjunto de dados alcançou uma adaptação perfeita aos dados de treino, mas com uma menor capacidade de generalização. A precisão nas previsões de teste é reduzida em comparação com o primeiro conjunto de dados.

Figura 33 – Árvore de Decisão no Segundo *Dataset*



Fonte: Elaborada pelo autor

5.3.6 Naive Bayes

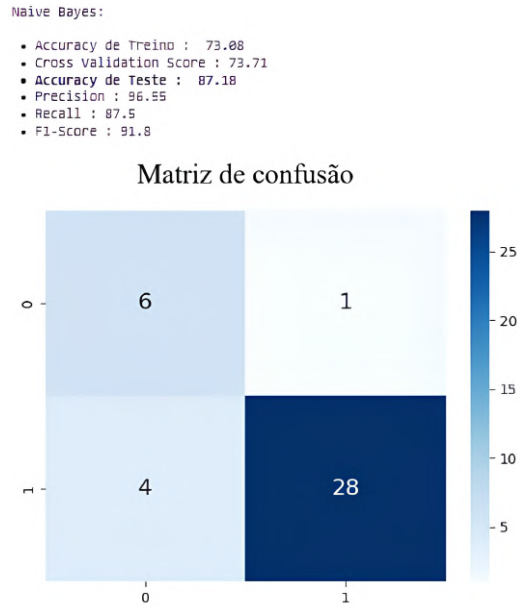
5.3.6.1 Primeiro Dataset

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 precisão de 96.55%, uma *recall* de 87.50% e um *F1-score* de 91.80%.

A acurácia de treino revela uma adaptação moderada aos dados de treino. O *score* de *cross validation* demonstra um desempenho consistente durante a validação cruzada. A acurácia de teste tem elevada precisão nas previsões em um contexto mais realista.

A precisão tem a maioria das previsões positivas corretas, indicando alto poder discriminativo. A *recall* tem identificação eficaz da maioria das instâncias positivas, minimizando falsos negativos. E o *F1-score* tem boa harmonia entre precisão e *recall*, indicando desempenho global sólido.

Os resultados do *Naive Bayes* no primeiro conjunto de dados mostram uma adaptação moderada aos dados de treino, mas um desempenho consistente durante a validação cruzada. A acurácia de teste é relativamente alta.

Figura 34 – *Naive Bayes* no Primeiro Dataset

Fonte: Elaborada pelo autor

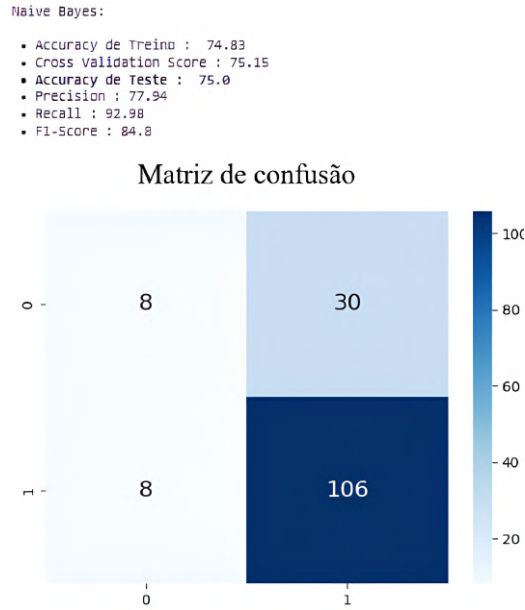
5.3.6.2 Segundo Dataset

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 precisão de 77.94%, uma *recall* de 92.98% e um F1-*score* de 84.80%.

A acurácia de treino mostra uma adaptação moderada aos dados de treino, porém menor em comparação com o primeiro conjunto. O *score* de *cross validation* indica consistência durante a validação cruzada, mas em um nível relativamente baixo. A acurácia de teste revela uma precisão moderada nas previsões em um contexto mais realista.

A precisão tem a maioria das previsões positivas corretas. A *recall* tem identificação eficaz da maioria das instâncias positivas, minimizando falsos negativos. E o F1-*score* apresenta um equilíbrio aceitável entre precisão e *recall*, refletindo um desempenho global razoável.

Os resultados do *Naive Bayes* no segundo conjunto de dados mostram uma adaptação moderada aos dados de treino, consistência durante a validação cruzada, mas um desempenho relativamente baixo em termos de acurácia de teste e *score* de *cross validation*.

Figura 35 – *Naive Bayes* no Segundo *Dataset*

Fonte: Elaborada pelo autor

5.3.7 XGBoost

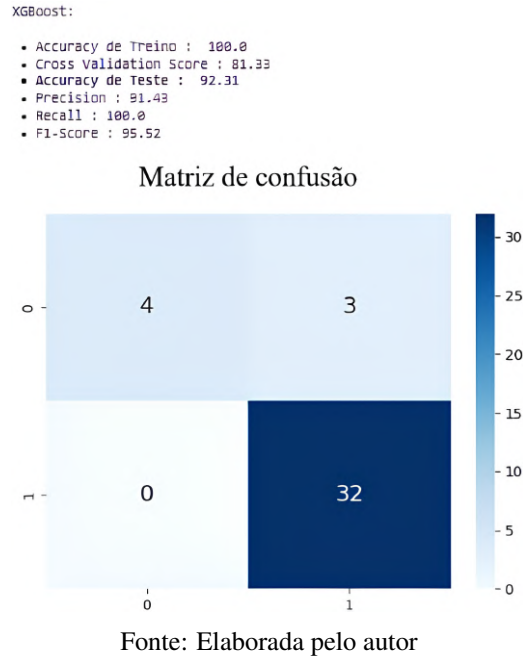
5.3.7.1 Primeiro Dataset

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 precisão de 91.43%, uma *recall* de 100.00% e um F1-*score* de 95.52%.

A acurácia de treino tem uma adaptação perfeita aos dados de treino, podendo sugerir *overfitting*. O *score* de *cross validation* demonstrou uma consistência moderada durante a validação cruzada, mas significativamente menor do que a acurácia de treino. A acurácia de teste apresenta uma alta precisão nas previsões para dados não vistos durante o treinamento.

A precisão indica que a maioria das previsões positivas é correta, sugerindo um poder discriminativo sólido. A *recall* tem identificação perfeita de todas as instâncias positivas. E o F1-*score* apresenta um equilíbrio excelente entre precisão e *recall*, indicando um desempenho global excepcional.

Os resultados do *XGBoost* no primeiro conjunto de dados mostram uma adaptação perfeita aos dados de treino, o que pode sugerir *overfitting*. No entanto, a acurácia de teste elevada e o equilíbrio entre precisão e *recall* indicam um bom desempenho geral na detecção da condição alvo.

Figura 36 – *XGBoost* no Primeiro *Dataset*

5.3.7.2 Segundo Dataset

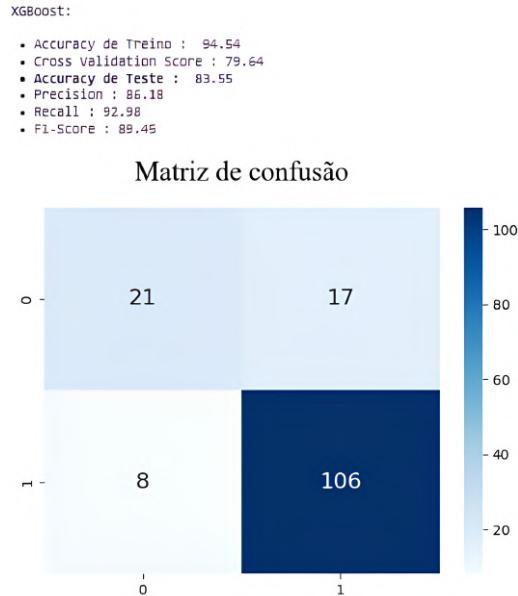
O *XGBoost* no segundo teve a acurácia de treino de 94.54%, um *score* de *cross validation* de 79.64%, uma acurácia de teste de 83.55%, uma precisão de 86.18%, uma *recall* de 92.98% e um *F1-score* de 89.45%.

A acurácia de treino indica uma adaptação eficaz aos dados de treino, evitando excesso de ajuste observado no primeiro conjunto de dados. O *score* de *cross validation* mostra uma consistência moderada durante a validação cruzada, com uma redução em comparação com a acurácia de treino. A acurácia de teste apresenta uma boa precisão nas previsões para dados não vistos durante o treinamento, embora ligeiramente inferior ao primeiro conjunto.

A precisão indica uma proporção significativa de previsões positivas corretas, contribuindo para uma boa capacidade discriminativa. O *recall* demonstra uma identificação eficiente da grande maioria das instâncias positivas. E o *F1-score* mostra um equilíbrio robusto entre precisão e *recall*, indicando um desempenho sólido na detecção da condição alvo.

Os resultados do *XGBoost* no segundo conjunto de dados são consistentes e eficazes, mantendo uma boa adaptação aos dados de treino, com destaque para a capacidade de generalização demonstrada pela acurácia de teste e equilíbrio entre precisão e *recall*.

Figura 37 – XGBoost no Segundo Dataset



Fonte: Elaborada pelo autor

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 foi treinado 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.

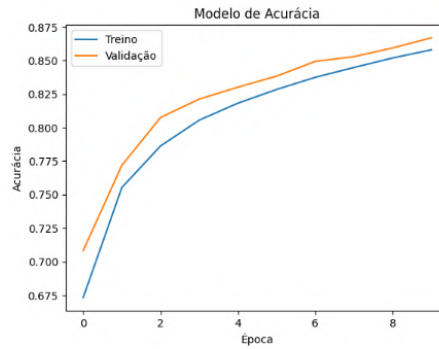
Figura 38 – Treinamento CNN

```
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
```

Fonte: Elaborada pelo autor

Pôde-se 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.

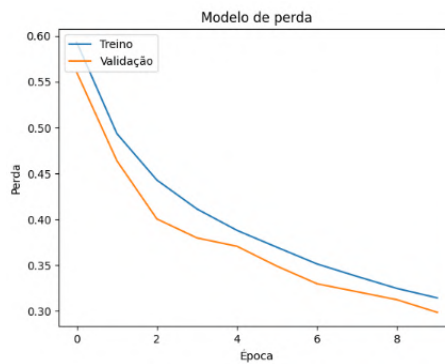
Figura 39 – Acurácia do Treinamento CNN



Fonte: Elaborada pelo autor

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

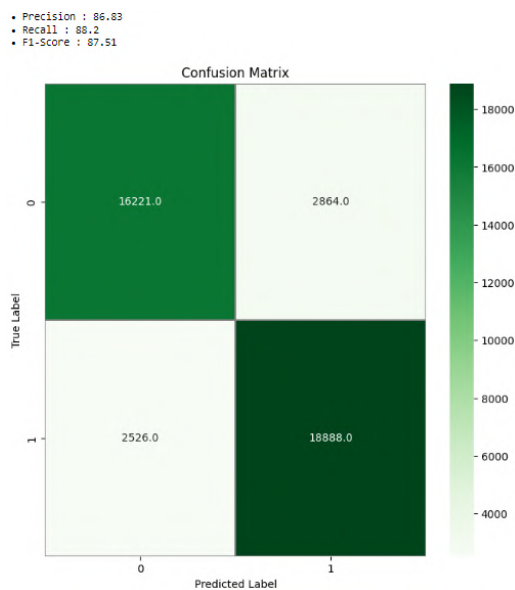
Figura 40 – Perda do Treinamento CNN



Fonte: Elaborada pelo autor

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

Figura 41 – Matriz de Confusão CNN



Fonte: Elaborada pelo autor

A configuração da rede neural de convolução ficou de acordo com a figura Figura 42.

Figura 42 – Configuração da Rede de Convolução

```
def get_conv_model():
    model = Sequential()
    model.add(Conv2D(16, (3, 3), activation='relu', strides = (1,1), padding = 'same', input_shape = input_shape))
    model.add(Conv2D(32, (3,3), activation='relu', strides = (1,1), padding = 'same', input_shape = input_shape))
    model.add(Conv2D(64, (3,3), activation='relu', strides = (1,1), padding = 'same', input_shape = input_shape))
    model.add(Conv2D(128, (3,3), activation='relu', strides = (1,1), padding = 'same', input_shape = input_shape))
    model.add(MaxPool2D((2,2)))
    model.add(Dropout(0.5))
    model.add(Flatten())
    model.add(Dense(128, activation='relu'))
    model.add(Dense(64, activation='relu'))
    model.add(Dense(10, activation='softmax'))
    model.summary()
    model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['acc'])
    return model
```

Fonte: Elaborada pelo autor

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, foi obtido uma acurácia de 0.7884 e uma perda de validação de 0.4325.

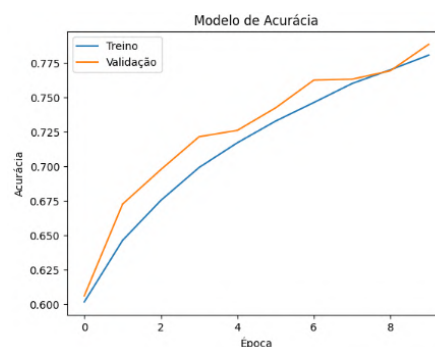
Figura 43 – Treinamento RNN

```
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
```

Fonte: Elaborada pelo autor

Ao analisar 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.

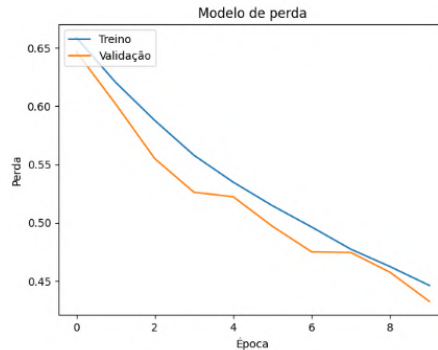
Figura 44 – Acurácia do Treinamento RNN



Fonte: Elaborada pelo autor

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.

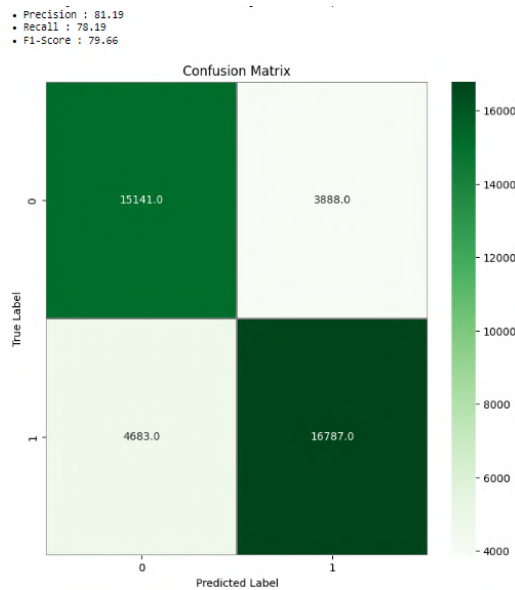
Figura 45 – Perda do Treinamento RNN



Fonte: Elaborada pelo autor

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

Figura 46 – Matriz de Confusão RNN



Fonte: Elaborada pelo autor

A configuração da rede neural recorrente ficou de acordo com a figura Figura 47.

Figura 47 – Configuração da Rede Recorrente

```

def get_recurrent_model():
    model = Sequential()
    model.add(LSTM(128, return_sequences=True, input_shape = input_shape))
    model.add(LSTM(128, return_sequences=True))
    model.add(Dropout(0.5))
    model.add(TimeDistributed(Dense(64, activation='relu')))
    model.add(TimeDistributed(Dense(32, activation='relu')))
    model.add(TimeDistributed(Dense(16, activation='relu')))
    model.add(TimeDistributed(Dense(8, activation='relu')))
    model.add(Flatten())
    model.add(Dense(10, activation='softmax'))
    model.summary()
    model.compile(loss='categorical_crossentropy', optimizer='adam', metrics=['acc'])
    return model

```

Fonte: Elaborada pelo autor

5.4 RESULTADOS

Ao realizar a análise dos resultados dos treinamentos dos modelos, foi obtido o seguinte para o primeiro dataset:

Quadro 2 – Resultados do primeiro *dataset*

Modelos	Acurácia	F1-Score
Regressão Logística	92.3076923076923	95.5223880597015
Support Vector Machine	92.3076923076923	95.5223880597015
K-Neighbors Classifier	89.74358974358975	93.93939393939394
Random Forest	92.3076923076923	95.5223880597015
Decision Tree	89.74358974358975	93.93939393939394
Naive Bayes	87.17948717948716	91.80327868852459
XGBoost	92.3076923076923	95.5223880597015

Fonte: Elaborada pelo autor

É visto 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%, considerando o F1-Score como critério de escolha os modelos são os mesmos porém com 95.52%.

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

Quadro 3 – Resultados do segundo *dataset*

Modelos	Acurácia	F1-Score
Regressão Logística	86.8421052631579	91.73553719008264
Support Vector Machine	84.21052631578947	90.2439024390244
K-Neighbors Classifier	82.23684210526315	88.20960698689956
Random Forest	82.23684210526315	87.3949579831933
Decision Tree	79.60526315789474	86.72566371681415
Naive Bayes	75.00000000000000	84.79999999999998
XGBoost	83.55263157894737	88.03418803418803

Fonte: Elaborada pelo autor

Consegue-se 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%, já com o *F1-Score* o melhor continuou a Regressão com 91.74%.

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.

Quadro 4 – Resultados do terceiro *dataset*

Modelos	Acurácia
Convolutional Neural Network	81.53
Recurrent Neural Network	76.70

Fonte: Elaborada pelo autor

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.

Dado que os resultados obtidos foram altos e satisfatórios, o objetivo de desenvolver um modelo de aprendizado de máquina para classificar pessoas com Parkinson e pessoas saudáveis, além de demonstrar a viabilidade em uso clínico, foi satisfeito

Para o trabalho futuro, visando ampliar a robustez e a aplicabilidade dos modelos de aprendizado de máquina no contexto da detecção de Parkinson temos que considerar a incorporação de modelos baseados em transformers, como BERT e GPT, na análise de dados relacionados à doença de Parkinson. Esses modelos, amplamente utilizados em processamento de linguagem natural, podem trazer benefícios significativos para a compreensão e extração de padrões complexos presentes em dados clínicos e de fala, a ampliação do número de épocas nas CNN e RNN para ver o impacto do aumento no número de épocas de treinamento nas redes e a expansão para mais datasets para testar e validar os modelos propostos em uma variedade mais ampla de conjuntos de dados relacionados à doença de Parkinson. Isso proporcionaria uma validação mais consistente e permitiria a generalização dos resultados para diferentes contextos clínicos.

REFERÊNCIAS

- ACADEMY, D. S. **Redes Neurais Recorrentes**. 2022. Disponível em: <<https://www.deeplearningbook.com.br/redes-neurais-recorrentes/>>. Acesso em: 19/08/2023.
- ACADEMY, D. S. **Redes Neurais Recorrentes**. 2022. Disponível em: <<https://www.deeplearningbook.com.br/algorithmo-backpropagation-parte-2-treinamento-de-redes-neurais/>>. Acesso em: 19/08/2023.
- 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.
- GEEKSFORGEEEKS. **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 NEURAIIS 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\(RNN,alimenta%20o%20apenas%20para%20camadas%20subsequentes\).>](https://imasters.com.br/data/um-mergulho-profundo-nas-redes-neurais-recorrentes#:~:text=Uma%20rede%20neural%20recorrente%20(RNN,alimenta%20o%20apenas%20para%20camadas%20subsequentes).>)>. Acesso em: 19/08/2023.

KARIMI, A. **Using a Hard Margin vs. Soft.** 2022. Disponível em: <<https://www.baeldung.com/cs/svm-hard-margin-vs-soft-margin>>. 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%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%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*, 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.