

Centro Universitário FEI



## **Relatório**

Classificação EEG Saudável vs. Epilético

PEL 301 - Biopotenciais

**Matheus Rodrigues Teixeira**

n° 124104-1

São Bernardo do Campo

2024

# 1 Objetivos

O objetivo deste estudo é avaliar a eficácia de diferentes modelos de aprendizado de máquina na classificação de sinais de EEG associados a crises epiléticas. Em particular, foi comparado o desempenho dos modelos Random Forest e Gradient Boosting utilizando características extraídas de sinais de EEG.

## 2 Introdução

A análise de sinais de EEG é essencial para o diagnóstico de condições neurológicas, como epilepsia. O reconhecimento automático de padrões em sinais de EEG pode melhorar a precisão do diagnóstico e reduzir o tempo necessário para a análise. Este relatório discute a aplicação de técnicas de aprendizado de máquina para classificar sinais de EEG em dois grupos, utilizando diferentes métodos de extração de características.

## 3 Metodologia

Os dados utilizados neste estudo são provenientes do conjunto de dados "Epileptic Seizure Recognition". A análise foi conduzida em três etapas principais:

- **Pré-processamento:** Remoção de ruído e ajuste do nível DC dos sinais de EEG.
- **Extração de Características:** Utilização de métodos de potência de EEG e transformada discreta de wavelet para extrair características dos sinais.
- **Classificação:** Aplicação de modelos Random Forest e Gradient Boosting para classificar os sinais utilizando as características extraídas.

Os dados foram inicialmente filtrados para remover os labels 2, 3 e 4, focando apenas nos labels de interesse (1 e 5). Posteriormente, foi realizado o ajuste do nível DC nos dados, seguido da geração de gráficos dos sinais para visualização. Foram utilizadas diferentes técnicas para extração de características, incluindo o uso de bandas de frequência do EEG e coeficientes da Transformada Discreta de Wavelet (TDW).

## 4 Resultados

Os modelos treinados utilizando características extraídas apresentaram um desempenho variável. O modelo Gradient Boosting alcançou a maior acurácia com 84,02%, enquanto o Random Forest atingiu 82,61%.

Tabela 1: Relatório de Classificação com Características do GradientBoosting-Classifier

Classe	Precisão	Recall	F1-Score	Suporte
1	86%	81%	83%	449
5	83%	87%	85%	471
<b>Acurácia</b>	84%			
<b>Média Macro</b>	84%	84%	84%	920
<b>Média Ponderada</b>	84%	84%	84%	920

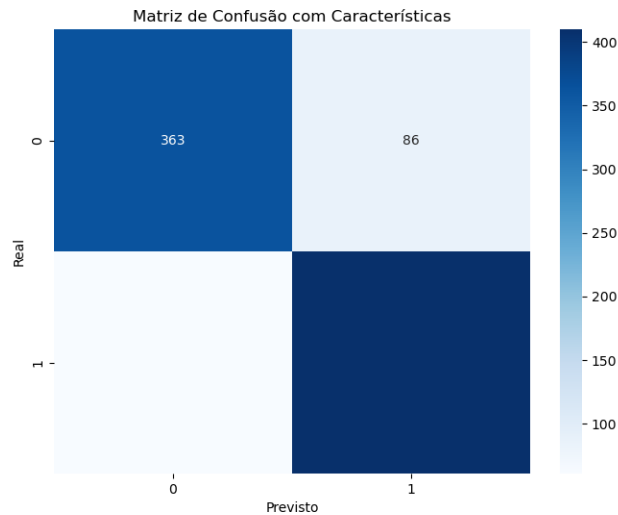


Figura 1: Matriz de Confusão para o Melhor Modelo com Características Extraídas

#### 4.1 Modelos com Características de Wavelet

A transformada discreta de wavelet foi utilizada para extrair coeficientes relevantes dos sinais de EEG. O modelo Gradient Boosting apresentou melhor desempenho comparado ao Random Forest.

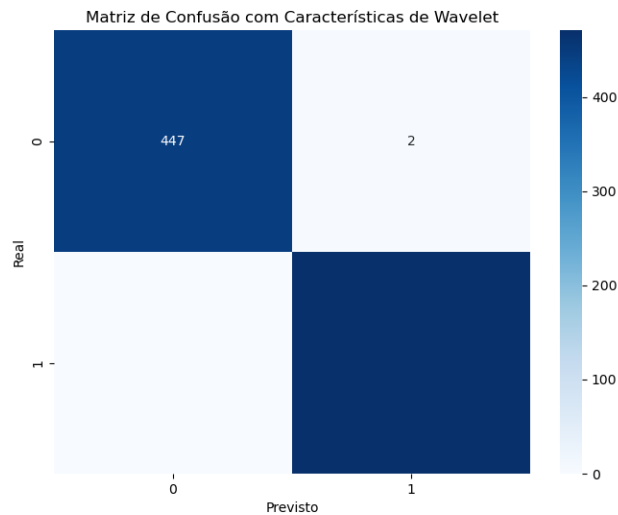


Figura 2: Matriz de Confusão para o Melhor Modelo com Características de Wavelet

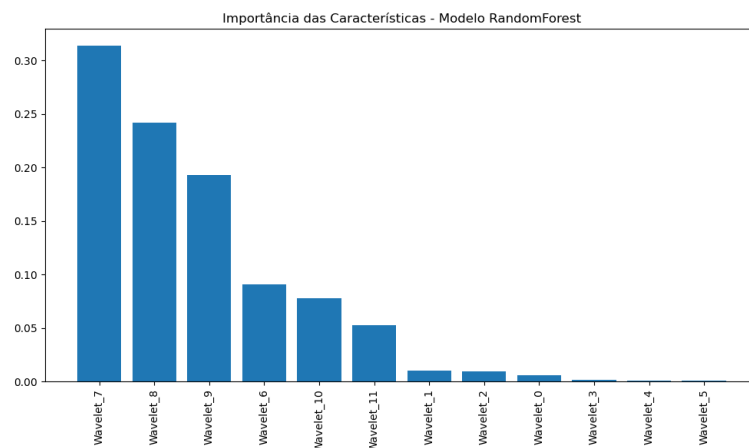
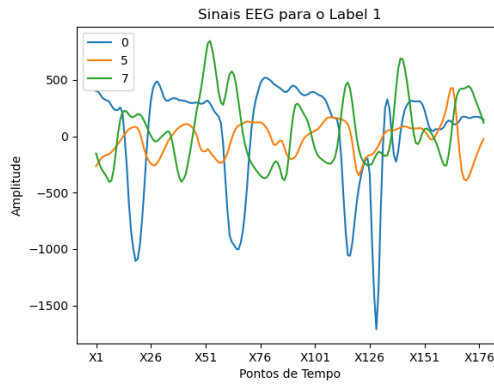


Figura 3: Importância das Características - Modelo RandomForest

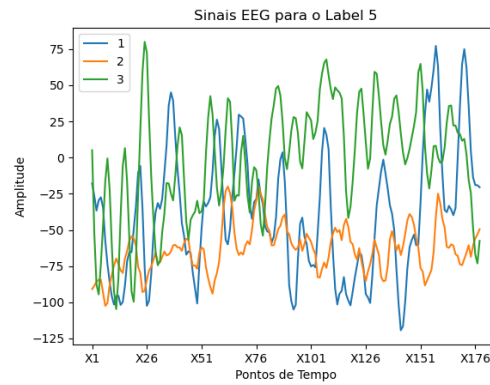
## 5 Conclusão

Os resultados indicam que o uso de características extraídas melhora significativamente o desempenho dos modelos de classificação de sinais de EEG. O modelo Gradient Boosting, em particular, apresentou a maior precisão em ambas as abordagens. No entanto, observou-se uma tendência ao overfitting no Random Forest quando utilizado com características de wavelet.

## Imagens Adicionais

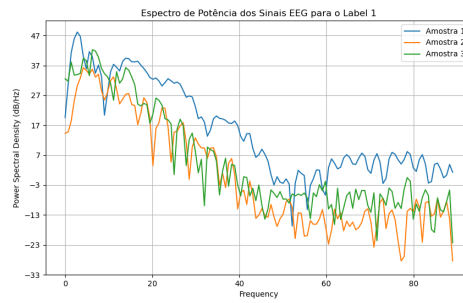


(a) Sinais EEG para Label 1

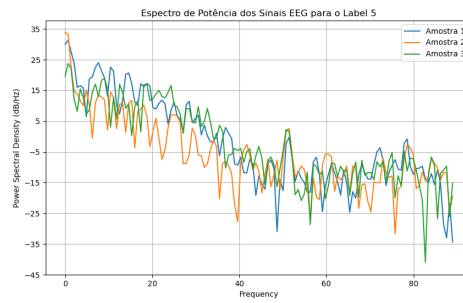


(b) Sinais EEG para Label 5

Figura 4: Exemplos de Sinais EEG por Label



(a) Espectro de Potência para Label 1



(b) Espectro de Potência para Label 5

Figura 5: Espectro de Potência dos Sinais EEG

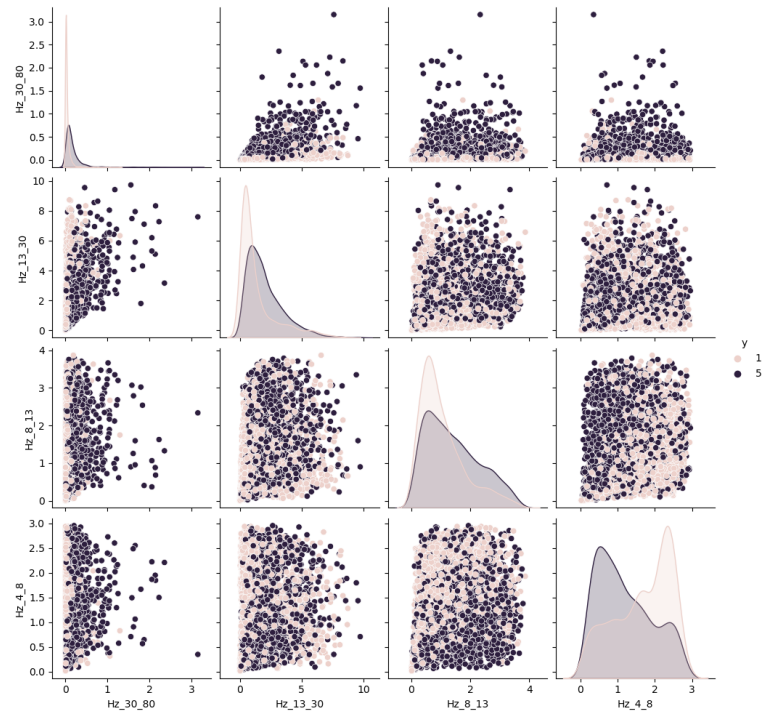


Figura 6: Distribuição de Características Extraídas dos Sinais EEG

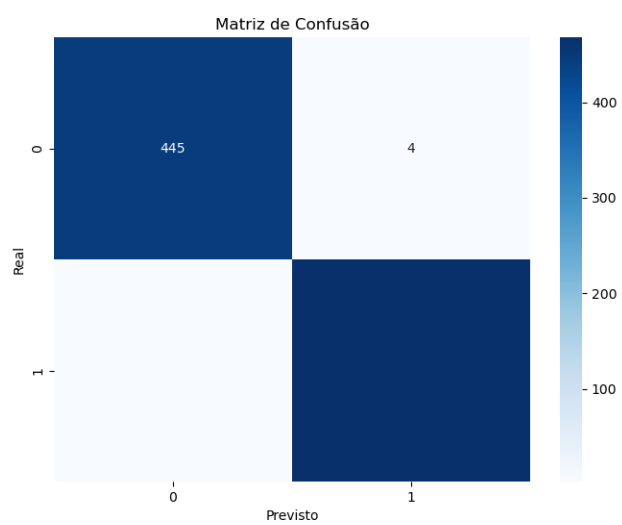


Figura 7: Matriz de Confusão do Melhor Modelo

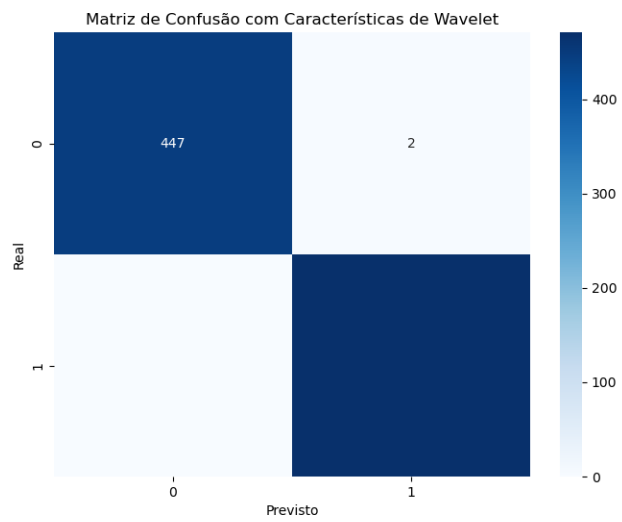


Figura 8: Matriz de Confusão com Características de Wavelet