

# Desenvolvimento de Modelos Computacionais para Análise de Eletroencefalograma e Detecção de Crises Epilépticas

Pedro Augusto Araujo da Silva de Almeida Nava Alves<sup>1</sup>,  
Alex Oliveira Barradas Filho<sup>1</sup>

<sup>1</sup>Universidade Federal do Maranhão (UFMA)  
65080-805 – São Luís – Maranhão – Brasil

pedro.alves@darti.ufma.br, alex.barradas@ecp.ufma.br

**Abstract.** *Epilepsy is a neurological disorder characterized by the recurrence of epileptic seizures. Epileptic seizures are associated with abnormal electrical activity in the brain. The electroencephalogram (EEG) is a monitoring method that records the electrical activity of the brain through electrodes placed on the scalp and is non-linear and dynamic in nature. Therefore, several studies have been done to develop alternative approaches for diagnosing epileptic seizures. It is observed in the literature that the application of machine learning techniques brings satisfactory results for the detection of epileptic seizures. Therefore, in this study, it was proposed to train several machine learning models with different preprocessing techniques to detect epileptic seizures, for comparative purposes. In the end, the best results were obtained with approximately 80% accuracy, for the Neural Networks and K Nearest Neighbors techniques using the Wavelet Transform for preprocessing, and for the Decision Tree technique using the Transform of Short Term Fourier.*

**Resumo.** *A epilepsia é um distúrbio neurológico caracterizado pela recorrência de crises epiléticas. As crises epiléticas estão associadas à atividade elétrica anormal no cérebro. O eletroencefalograma (EEG) é um método de monitoramento que registra a atividade elétrica do cérebro por meio de eletrodos colocados no couro cabeludo e é de natureza não linear e dinâmica. Portanto, vários estudos foram feitos para desenvolver abordagens alternativas para o diagnóstico de crises epiléticas. Observa-se na literatura que a aplicação de técnicas de aprendizado de máquina traz resultados satisfatórios para detecção de crises epiléticas. Portanto, neste estudo, foi proposto treinar diversos modelos de aprendizado de máquina com diferentes técnicas de pré-processamento para detecção de crises epiléticas, para efeitos comparativos. Ao fim, obteve-se os melhores resultados com aproximadamente 80% de acurácia, para as técnicas de Redes Neurais e K Vizinhos mais Próximos utilizando a Transformada Wavelet para o pré-processamento, e para a técnica de Arvore de Decisão usando a Transformada de Fourier de Curto Termo.*

## 1. Introdução

A epilepsia é uma doença neurológica causada por diferentes etiologias e caracterizada pela ocorrência de crises epiléticas recorrentes [10]. As crises epiléticas têm intensidade e intervalo variáveis, podendo causar episódios breves ou longos períodos de

convulsão resultando em lesões físicas. Essa condição prejudica diretamente a qualidade de vida, pois tem consequências neurológicas, cognitivas, psicológicas e sociais.

As crises epiléticas estão associadas a atividade elétrica anormal e excessiva no cérebro. O diagnóstico de crises epiléticas envolve consultar o histórico médico do paciente e identificar descargas ictais ou interictais (IED) no eletroencefalograma (EEG) [13].

O Eletroencefalograma (EEG) é um método de monitoramento eletrofisiológico não invasivo que registra a atividade elétrica do cérebro colocando eletrodos em diferentes posições no couro cabeludo [11]. O EEG mede a atividade a partir das flutuações de voltagem dos eletrodos ao longo do tempo que representam a atividade cerebral naquela região. O EEG é considerado um método padrão para detecção de epilepsia devido a sua ampla disponibilidade e baixo custo. Porém, o EEG apresenta informações complexas do paciente, que necessitam de análises longas e passíveis de erros quando realizadas manualmente.

Este cenário tem motivado o aumento de abordagens alternativas para o diagnóstico de crises epiléticas [6, 1]. Esses métodos extraem características lineares e não lineares no domínio da frequência ou do tempo dos sinais de EEG. Em seguida, esses recursos extraídos são usados a partir de técnicas de aprendizado de máquina para detecção automática de crises epiléticas [17, 2].

Diante disso, o objetivo deste trabalho é analisar os principais métodos de análise de EEG no domínio da frequência, e em seguida, utilizar diferentes técnicas de aprendizagem de máquina, para classificação de crises epiléticas, para efeitos comparativos.

## 2. Metodologia

A metodologia proposta neste trabalho está ilustrada pela Figura 1. Ela consiste em: Aquisição da base de dados com os eletroencefalogramas; Pré-processamento da base de dados por técnicas de análise no domínio da frequência; Extração das características nas base de dados pré-processada; aplicação de técnicas de aprendizagem de máquina para detecção de crises epiléticas; Análise dos resultados obtidos.

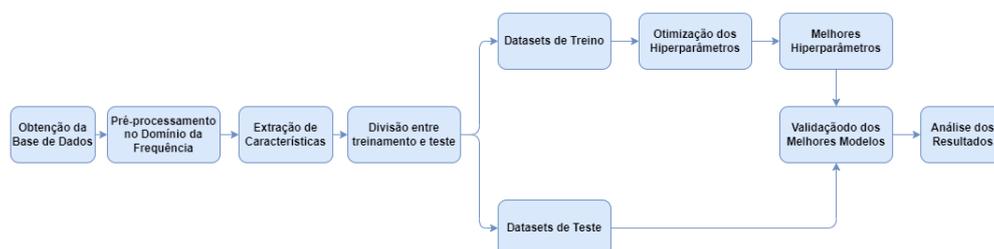


Figura 1. Metodologia Proposta

### 2.1. Banco de Dados

O banco de dados foi obtido a partir dos eletroencefalogramas coletados e disponibilizados pelo *Temple University Hospital (TUH)*. O banco de dados disponível contém 1947 registros, sendo 571 classificados com crises epiléticas e 1376 classificados sem crises epiléticas. Os sinais do eletroencefalograma foram coletados por meio

de eletrodos colocados no couro cabeludo dos pacientes a uma taxa de amostragem de 250Hz. Foi utilizado o *International System 10-20*, posicionando os 2 eletrodos na região pré-frontal (Fp); 5 eletrodos na região frontal (F); 4 eletrodos na região temporal (T); 3 eletrodos na região parietal (P); 3 eletrodos na região central (C); e 2 eletrodos na região occipital (O). Os sinais do eletroencefalograma ficam então disponibilizados no formato EDF (*European Data Format*) com as gravações dos 19 canais do eletroencefalograma realizado.

## 2.2. Pré-processamento

O banco de dados foi pré-processado para extrair as características utilizadas no treinamento do modelo. As ondas cerebrais *delta*, *theta*, *alpha*, *beta* e *gamma* foram extraídas de cada canal do eletroencefalograma [5, 16]. Um filtro de passagem de banda foi aplicado nos intervalos: 0,5 Hz a 4 Hz, 4 Hz a 8 Hz, 8 Hz a 12 Hz, 12 Hz a 35 Hz e 35 Hz a 100 Hz. Posteriormente, cinco técnicas de análise no domínio da frequência foram aplicadas: Transformada de Fourier [9], Transformada de Cosseno [3], Transformada de Fourier de Curto Termo [4], Transformada Wavelet [14] e Decomposição de Pacotes Wavelet [20]. As técnicas foram escolhidas de acordo com os trabalhos realizados na literatura [1, 6, 2, 13, 19]

### 2.2.1. Transformada de Fourier

A transformada de Fourier é uma operação que decompõe uma função temporal  $f(t)$  no domínio da frequência usando a seguinte equação:

$$F(\xi) = \int_{-\infty}^{\infty} f(t) \cdot e^{-2\pi i \xi t} dt \quad (1)$$

O resultado dessa operação é uma função complexa  $F(\xi)$  cujo valor absoluto representa a quantidade de frequência  $\xi$  presente na função original. Para sinais discretos, como os sinais digitais do eletroencefalograma, é aplicada a transformada de Fourier discreta, representada pela seguinte equação:

$$X_k = \sum_{n=0}^{N-1} x[n] \cdot e^{-\frac{2\pi i}{N} kn} \quad (2)$$

O resultado é um vetor  $\{X_k\} := X_0, X_1, \dots, X_{N-1}$  com cada posição sendo um número complexo que codifica amplitude e fase do componente sinusoidal que representa o domínio da frequência.

### 2.2.2. Transformada de Cosseno

Na transformada Cosseno, é aplicada uma variação da transformada de Fourier, na qual apenas o cosseno da fórmula de Euler é usado, definido pela equação:

$$F(\nu) = \int_{-\infty}^{\infty} f(t) \cdot \cos(-2\pi \nu t) dt \quad (3)$$

A transformada discreta de cosseno é definida igual à transformada discreta de Fourier. Porém, Ela obtém apenas componentes reais do domínio da frequência, então é computacionalmente mais eficiente.

### 2.2.3. Transformada de Fourier de Curto Termo

A Transformada de Fourier de Curto Termo (STF) é uma derivação não estacionária da Transformada de Fourier. O espectro de frequência obtido por esta transformada varia com o tempo e, portanto, é não estacionário. Essa técnica consiste em aplicar a transformada de Fourier em diferentes momentos do sinal original, chamados de termos curtos, que são obtidos usando uma função de janelamento  $g(\tau)$ . Assim, a equação incluindo a função de janelamento torna-se:

$$F(\xi, \tau) = \int_{-\infty}^{\infty} f(t)g(t - \tau) \cdot e^{-2\pi i \xi t} dt \quad (4)$$

O resultado da transformada é o espectro de frequência obtido pela transformada de Fourier, em um curto intervalo de tempo da função original definida pela função de janelamento  $g(\tau)$ .

Para sinais digitais, a versão discreta da Transformada de Fourier de curto prazo também é usada. A transformada discreta de Fourier também é aplicada ao sinal digital em diferentes termos curtos do sinal, usando uma função de janela discreta  $h[t]$ . Assim, A equação da Transformação Discreta de Fourier de Curto Termo torna-se:

$$X_{k,m} = \sum_{n=0}^{N-1} x[n]h[n - m] \cdot e^{-\frac{2\pi i}{N} kn} \quad (5)$$

### 2.2.4. Transformada Wavelet

A transformada wavelet é definida pelo produto interno entre a função de entrada e a função wavelet após um fator de escala  $a$  e fator de deslocamento  $b$ , denotado pela seguinte equação:

$$F(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{\psi\left(\frac{x - b}{a}\right)} \cdot f(x) dx \quad (6)$$

Os fatores de escala e deslocamento ( $a$  e  $b$ ) permitem obter informações sobre o sinal original em altas e baixas frequências e em diferentes momentos do sinal. A função wavelet  $\psi$  é usada como um impulso aplicado ao sistema representado pelo sinal original [14].

A transformada wavelet discreta é usada para sinais digitais. Na transformada wavelet discreta, o fator de escala  $a$  é definido em valores discretos de  $a_j = 2^j$ . O fator de deslocamento também é definido em valores discretos de  $b_j = 2^j k$ . O algoritmo da transformada wavelet discreta divide o processo em  $n = \log_2(N)$  níveis. Em cada nível

$j$ , um filtro passa-baixa e passa-alta é aplicado no sinal de saída  $y_{j-1}[k]$  do nível anterior, de acordo com as equações 7 e 8

$$y_{low} = \sum_{k=-\infty}^{\infty} x[k] \cdot g[2n - k] \quad (7)$$

$$y_{high} = \sum_{k=-\infty}^{\infty} x[k] \cdot h[2n - k] \quad (8)$$

A  $h[t]$  é a função de impulso do filtro passa-alta definido pela função wavelet com valores discretos, representado pela equação 9, e  $g[t]$  seu filtro de espelho de quadratura. Os coeficientes discretos  $y_{high}$ , também chamados de componentes detalhados, têm apenas metade das frequências para o nível  $j$  e são a saída da transformada wavelet para o nível  $j$ . Os coeficientes discretos  $y_{low}$ , também chamados de componentes aproximados, têm a outra metade das frequências. Os componentes são então passados recursivamente para o algoritmo do próximo nível  $j + 1$ . Assim, a transformada calcula os componentes detalhados e aproximados em cada nível e passa os componentes aproximados para calcular os componentes do próximo nível. A transformada Wavelet discreta final é formada pelo conjunto de todos os componentes detalhados gerados em cada nível com o componente aproximado gerado no último nível, resultando em  $n + 1$  coeficientes.

$$h[t] = \frac{1}{\sqrt{2^j}} \cdot \psi\left(\frac{-t}{2^j}\right), t = 1, 2^j, 2^{2j}, \dots, 2^N \quad (9)$$

### 2.2.5. Decomposição por Pacotes Wavelet

Decomposição por Pacotes Wavelet é uma variação da Transformada Wavelet Discreta onde mais filtros são aplicados ao sinal no algoritmo para calcular os coeficientes. Na Decomposição por Pacotes Wavelet, ambos componentes detalhados e aproximados são passados para o próximo nível. Portanto, o algoritmo cresce exponencialmente com  $2^n$  coeficientes, enquanto que a transformada Wavelet tem  $n + 1$  coeficientes.

### 2.3. Extração de Características

Após extrair os cinco intervalos de ondas cerebrais para cada um dos 19 canais, 95 ondas foram obtidas para cada eletroencefalograma. Em seguida, o método de pré-processamento foi aplicado nos sinais extraídos para cada um dos cinco métodos de pré-processamento. Assim, cinco conjuntos de dados diferentes foram obtidos para a análise no domínio da frequência, um para cada método. As características do conjunto de dados foram extraídas calculando-se cinco características estatísticas: valor médio, desvio padrão, energia média, valor absoluto médio, assimetria. Essas características são definidas respectivamente nas seguintes equações:

$$\mu = \frac{1}{N} \sum_{k=0}^{N-1} x[k] \quad (10)$$

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{k=0}^{N-1} (x[k] - \mu)^2} \quad (11)$$

$$AE = \frac{1}{N} \sum_{k=0}^{N-1} (x[k])^2 \quad (12)$$

$$AAV = \frac{1}{N} \sum_{k=0}^{N-1} |x[k]| \quad (13)$$

$$A = \frac{\sum_{k=0}^{N-1} (x[k] - \mu)^3}{N * \sigma^3} \quad (14)$$

## 2.4. Aprendizagem de Máquina

O banco de dados possui um total de 1142 amostras. O conjunto de dados foi dividido em treinamento e teste, de forma balanceada para evitar um *bias* durante a seleção do modelo mais otimizado. 80% do conjunto de dados foi usado para treinamento e 20% para teste. O conjunto de dados de treinamento foi usado na otimização do hiperparâmetro e o de teste foi usado para avaliar o melhor modelo entre as técnicas.

As técnicas de aprendizagem de máquina utilizadas foram: Máquina de Vetores de Suporte (SVM) [15], Regressão Logística, Árvores de Decisão [18], K Vizinhos Mais Próximo (KNN) [7], Redes Neurais: Perceptron de Multicamadas (MLP) [12] e Otimização por Gradiente Descendente Estocástico (SGD) [8]. Cada método de aprendizado de máquina foi otimizado ajustando seus hiperparâmetros através do método *grid search*. Esse método consiste em determinar as possíveis combinações dos valores de hiperparâmetros, e selecionar a melhor combinação de hiperparâmetros ao avaliar os modelos treinados para cada combinação determinada. Para selecionar o melhor hiperparâmetro durante a otimização, a validação cruzada *K Fold* [15] foi usada com  $k = 10$ . O *K Fold* consiste em dividir ainda mais o conjunto de dados de treinamento em  $k$  partições, com  $k - 1$  usadas para treinamento e 1 para teste. A partição usada para teste é permutada de forma que todas as  $k$  partições sejam usadas pelo menos 1 vez para teste. A avaliação da combinação de hiperparâmetros é a média das métricas dos  $k$  modelos treinados. Para a melhor seleção do hiperparâmetro, foi utilizada a métrica de acurácia, definida pela equação 17.

Após a otimização, os melhores valores de hiperparâmetros foram selecionados para cada método de aprendizado de máquina e conjunto de dados de pré-processamento. Em seguida, um modelo foi treinado com os valores dos hiperparâmetros em todo o conjunto de dados de treinamento. Por fim, o melhor modelo de cada técnica foi avaliado no conjunto de dados de teste, usando as métricas de classificação: precisão, sensibilidade, acurácia e Teste F1. As métricas de classificação são definidas pelas equações 15, 16, 17 e 18 respectivamente. Cada métrica usa a matriz de confusão calculada. Os valores de TP, TN, FN e FP usados na equação representam os valores verdadeiro-positivo, verdadeiro-negativo, falso-negativo e falso-positivo, respectivamente, da matriz de confusão.

$$precisão = \frac{TP}{TP + FP} \quad (15)$$

$$\text{sensibilidade} = \frac{TP}{TP + FN} \quad (16)$$

$$\text{acurácia} = \frac{TP + TN}{TP + FP + TN + FN} \quad (17)$$

$$F_1 = \frac{2TP}{2TP + FP + FN} \quad (18)$$

### 3. Resultados

Cada método de aprendizagem de máquina foi aplicado em cada base de dados de pré-processamento. Os hiperparâmetros foram otimizados com a métrica de precisão usada para seleção dos melhores hiperparâmetros. Os resultados da otimização para cada método de aprendizagem de máquina por base de dados são representados pelas Tabelas 1, 2, 3, 4, 5 e 6. Cada linha apresenta os 3 melhores modelos para cada coluna do método de pré-processamento, com os melhores resultados em negrito.

Os resultados para a SVM são próximos, independentemente da base de dados pré-processada utilizada. Os resultados do KNN, Regressão Logística e MLP têm um aumento entre 3% e 5% para as base de dados pré-processadas com a Transformada Wavelet e Decomposição por Pacotes Wavelet. Os resultados da Otimização por Gradiente Descendente Estocástico foram maiores para a Transformada de Fourier e Transformada de Fourier de Curto Termo, enquanto que os resultados para a Árvore de Decisão foram maiores apenas para a Transformada de Fourier de Curto Termo. Os métodos que obtiveram os melhores resultados foram a Árvores de Decisão, MLP e KNN, com melhores valores próximos a 80%. A Regressão Logística também obteve resultados próximos, com melhores valores em torno de 7%. Tanto a SVM quanto a Otimização por Gradiente Descendente Estocástico obtiveram os piores resultados, com melhores valores próximos a 55%.

Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
77.0831%	77.4104%	76.4226%	<b>78.9465%</b>	<b>80.2628%</b>
73.9047%	77.1859%	74.3442%	78.1808%	79.6046%
73.5702%	76.9709%	73.9142%	78.1713%	79.0612%

**Tabela 1. Otimização do KNN**

Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
<b>51.4238%</b>	<b>51.8645%</b>	<b>51.2076%</b>	<b>51.9709%</b>	<b>51.7559%</b>
51.4238%	51.757%	51.0977%	51.7547%	51.0977%
51.3151%	51.4262%	50.989%	51.5337%	51.0977%

**Tabela 2. Otimização da SVM**

O melhor modelo de cada método de aprendizagem de máquina e cada base de dados pré-processada foram avaliados usando as outras métricas de classificação no conjunto de dados de teste e seus resultados são representados pela Figura 2. O KNN obteve

Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
74.6751%	76.4178%	75.7561%	<b>79.5975%</b>	<b>78.5069%</b>
74.447%	76.0882%	74.232%	79.2726%	77.9443%
74.3383%	75.7561%	74.0002%	78.397%	77.8488%

**Tabela 3. Otimização do MLP**

Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
68.0948%	66.9828%	67.8655%	<b>70.0669%</b>	<b>71.8156%</b>
68.0948%	66.8813%	67.7556%	69.0815%	71.1538%
67.7676%	66.8765%	67.3184%	69.0815%	70.8242%

**Tabela 4. Otimização da Regressão Logística**

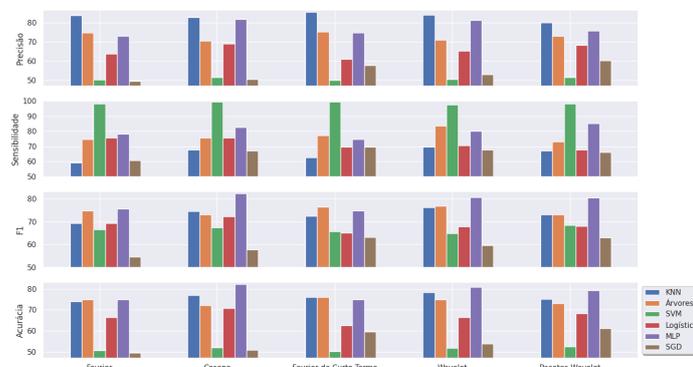
Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
<b>59.9737%</b>	52.629%	<b>58.7733%</b>	56.5815%	53.4245%
59.7611%	51.9709%	57.1249%	55.2699%	53.4161%
59.1042%	50.6605%	55.1517%	54.0576%	52.4283%

**Tabela 5. Otimização do SGD**

Fourier	Cosseno	STF	Wavelet	Pacotes Wavelet
75.9974%	75%	<b>79.3968%</b>	73.5738%	74.4505%
75.6677%	74.8901%	78.7339%	73.2465%	74.2284%
75.6677%	74.6715%	78.5165%	73.1378%	74.1245%

**Tabela 6. Otimização da Árvores de Decisão**

o melhor resultado para a métrica de precisão, mas tiveram resultados baixos para a sensibilidade. A SVM teve os melhores valores para a sensibilidade, mas os piores valores para as outras métricas. Os modelos com melhores resultados para todas as métricas foi a MLP para os métodos de pré-processamento Transformada de Cosseno e Transformada Wavelet. Ambos os modelos têm pelo menos 80% para todas as 4 métricas.



**Figura 2. Resultados de todas as métricas**

#### 4. Conclusão

Este trabalho aplicou alguns dos principais métodos de técnicas de aprendizagem de máquina e análise no domínio da frequência para desenvolver um modelo classificador de crises epiléticas em eletroencefalograma. Cinco métodos diferentes de pré-processamento foram usados para comparar os resultados: Transformada de Fourier, Transformada Cosseno, Transformada de Fourier de Curto Prazo, Transformada Wavelet e Decomposição de Pacote Wavelet. Após a extração de suas características, seis métodos de Aprendizado de Máquina foram aplicados: K Vizinhos Mais Próximos, Máquinas de Vetores de Suporte, Árvores de Decisão, Regressão Logística, Otimização por Gradiente Descendente Estocástico e Rede Neural: Perceptron Multicamadas.

Apos obtidos os resultados dos melhores modelos desenvolvidos, pode-se observar que a técnica com os melhores resultados para a precisão foi o KNN. Isto indica que a técnica possui uma baixa taxa de falso-positivo, porém a baixa sensibilidade indica alta taxa de falso-negativo. Para a sensibilidade, a técnica com os melhores resultados foi o SVM, que possui as vantagens e desvantagens opostas ao KNN, com a baixa taxa de falso-negativo, mas alta taxa de falso-positivo devido ao baixos valores da precisão. Portanto para garantir que a técnica possua baixa taxa de falso-positivo e falso-negativo, as métricas F1 e Acurácia. Para ambas as métricas, a técnica MLP obteve os melhores resultados dentre as 5 técnicas. Apesar de que a MLP obteve os melhores resultados em geral, alguns dos outros modelos obtiveram resultados próximos para certas técnicas de pré-processamento, e portanto, para a escolha de um modelo final a ser utilizado em uma aplicação real é necessário alguma análise estatística posterior em cima da acurácia ou métrica F1 destes modelos de resultados próximos. Todavia, a partir da análise dos resultados entre as técnicas de aprendizagem de máquina e pré-processamento, pode-se concluir que dentre as técnicas analisadas, a MLP foi a técnica com os melhores resultados em geral, enquanto que as técnicas de Transformadas Wavelet e Cosseno obtiveram os melhores resultados entre os métodos de pré-processamento.

Apesar dos resultados satisfatórios da MLP, observa-se na literatura que é possível obter resultados melhores, utilizando modelos mais complexos, como redes convolucionais [6, 1, 11, 17, 2]. Porém, como visto na literatura, a técnica de pré-processamento pela Transformada Wavelet, é a que obtém os melhores resultados devida a alta complexidade de análise em altas e baixas frequências, ao longo do tempo, como foi obtido neste trabalho.

#### Referências

- [1] Acharya, U.R., Subbhuraam, V.S., Ang, P., Yanti, R., Suri, J.: Application of non-linear and wavelet based features for the automated identification of epileptic eeg signals. *International journal of neural systems* **22**, 1250002 (04 2012)
- [2] Adeli, H., Zhou, Z., Dadmehr, N.: Analysis of eeg records in an epileptic patient using wavelet transform. *Journal of Neuroscience Methods* **123** (2003)
- [3] Ahmed, N., Natarajan, T., Rao, K.: Discrete cosine transform. *IEEE Transactions on Computers* **C-23**(1), 90–93 (1974)
- [4] Allen, J.: Short term spectral analysis, synthesis, and modification by discrete fourier transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing* **25**(3), 235–238 (1977)

- [5] Amorim, E., van der Stoel, M., Nagaraj, S.B., Ghassemi, M.M., Jing, J., O'Reilly, U.M., Scirica, B.M., Lee, J.W., Cash, S.S., Westover, M.B.: Quantitative eeg reactivity and machine learning for prognostication in hypoxic-ischemic brain injury. *Clinical Neurophysiology* **130** (2019)
- [6] Bajaj, V., Pachori, R.B.: Classification of seizure and nonseizure eeg signals using empirical mode decomposition. *IEEE Transactions on Information Technology in Biomedicine* **16**(6), 1135–1142 (2012)
- [7] Batista, G.E.A.P.A., Silva, D.F.: How k-nearest neighbor parameters affect its performance (2009)
- [8] Bottou, L.: Large-scale machine learning with stochastic gradient descent. In: Lechevalier, Y., Saporta, G. (eds.) *Proceedings of COMPSTAT'2010*. pp. 177–186. Physica-Verlag HD, Heidelberg (2010)
- [9] Bracewell, R.N.: The fourier transform. *Scientific American* **260**(6), 86–95 (1989)
- [10] BS, C., DH, L.: Epilepsy. *The New England Journal of Medicine* **349** (2003)
- [11] Duan, L., Bao, M., Miao, J., Xu, Y., Chen, J.: Classification based on multilayer extreme learning machine for motor imagery task from eeg signals. *Procedia Computer Science* **88** (2016)
- [12] Gardner, M., Dorling, S.: Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. *Atmospheric Environment* **32**(14), 2627–2636 (1998)
- [13] Joshi, V., Pachori, R.B., Vijesh, A.: Classification of ictal and seizure-free eeg signals using fractional linear prediction. *Biomedical Signal Processing and Control* **9**, 1–5 (2014)
- [14] Junsheng, C., Dejie, Y., Yu, Y.: Application of an impulse response wavelet to fault diagnosis of rolling bearings. *Mechanical Systems and Signal Processing* **21** (2007)
- [15] Ling, H., Qian, C., Kang, W., Liang, C., Chen, H.: Combination of support vector machine and k-fold cross validation to predict compressive strength of concrete in marine environment. *Construction and Building Materials* **206**, 355–363 (2019)
- [16] Luo, J., Gao, X., Zhu, X., Wang, B., Lu, N., Wang, J.: Motor imagery eeg classification based on ensemble support vector learning. *Computer Methods and Programs in Biomedicine* **193** (2020)
- [17] Malmivuo, J., Plonsey, R.: *Bioelectromagnetism*. 13. Electroencephalography, pp. 247–264 (01 1995)
- [18] Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., Brown, S.D.: An introduction to decision tree modeling. *Journal of Chemometrics* **18**(6), 275–285 (2004)
- [19] Raghu, S., Sriraam, N., Temel, Y., Shyam Vasudeva Rao, P.L.K.: Eeg based multi-class seizure type classification using convolutional neural network and transfer learning. *Neural Networks* **124** (2020)
- [20] Ting, W., Guo-zheng, Y., Bang-hua, Y., Hong, S.: Eeg feature extraction based on wavelet packet decomposition for brain computer interface. *Measurement* **41**(6), 618–625 (2008)