

RELATÓRIO PARCIAL - IFSP

Nome do Bolsista: Luís Otávio Lopes Amorim

Nome do Orientador: Dr. Ricardo Pires

Título do Projeto: Comparação entre métodos de escolha de janelas temporais para treino em sistema para previsão de surtos epilépticos

Vigência da Bolsa: 24/04/20 a 30/11/21

INTRODUÇÃO

A epilepsia é um dos distúrbios neurológicos mais comuns, afetando 65 milhões de pessoas no mundo (EPILEPSY FOUNDATION, 2021; WORLD HEALTH ORGANIZATION, 2019). Ela é diagnosticada após uma pessoa ter tido ao menos dois surtos sem nenhuma outra causa médica conhecida.

As convulsões sofridas pelos pacientes ocorrem devido a distúrbios na atividade elétrica do cérebro. Esses distúrbios podem estar relacionados a lesões no cérebro ou mesmo causas genéticas, ainda assim, na maioria dos casos a causa é desconhecida (SHAFER; SIRVEN, 2020). Surtos podem ser desde pequenos lapsos de atenção até convulsões prolongadas e variam, em frequência, desde menos de um por ano até vários por dia (WORLD HEALTH ORGANIZATION, 2019).

A atividade cerebral é medida na forma do chamado eletroencefalograma (EEG), usando-se eletrodos metálicos posicionados na superfície do crânio. O EEG manifesta-se, dessa forma, como ondas ao longo do tempo, uma para cada eletrodo (também chamados de canais), representando as sinapses das comunicações entre as células cerebrais. Esse exame é utilizado nos principais testes para diagnóstico de epilepsia, além de outros distúrbios neurológicos (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH, 2020).

Desta forma os períodos de atividade cerebral são definidos como ictal, correspondendo ao período de um surto; pré-ictal, que precede imediatamente um surto; pós-ictal, imediatamente a seguir de um surto e interictal, período distante de surtos (MULA; MONACO, 2011).

Os tratamentos para a doença são baseados em medicamentos ou cirurgia. Eles são ineficazes para a maioria dos pacientes (GADHOUMI et al., 2016). Dessa forma há grande interesse no desenvolvimento de sistemas para a previsão de surtos epilépticos para serem utilizados no dia a dia dos pacientes. Se um sistema desse tipo advertir um paciente sobre a possível chegada de um surto com uma boa antecedência e boa precisão, esse paciente pode colocar-se em situação de segurança, como parar de dirigir um carro. Além disso, para os pacientes para os quais os medicamentos são eficazes, eles podem ser administrados apenas após a advertência desse sistema, minimizando assim efeitos colaterais.

Tsiouris et al. realizaram um trabalho muito promissor utilizando uma rede neural do tipo Long Short Term Memory (LSTM) (HOCHREITER; SCHMIDHUBER, 1997). Por mais que os resultados apresentados

por esse grupo sejam promissores, a forma com que a separação do conjunto de treino e teste foi realizada é discutível já que, os vetores de uma única crise foram utilizados nos dois conjuntos de dados, assim pode-se dizer que a rede neural utilizou algumas informações que, em teoria, deveriam estar presentes apenas no grupo de teste, para ser treinada.

OBJETIVO ou PROPOSIÇÃO

O objetivo geral deste projeto é a realização de experimentos de treinamento e de teste de sistema de predição de surtos epilépticos com base no artigo de Tsiouris et al. (2018), variando-se a forma de escolha das janelas temporais para treino conforme descrito na introdução, para avaliação de seu impacto na acurácia dos resultados obtidos.

Os objetivos específicos são:

- Compreender o processo de extração de características de EEG usado por Tsiouris et al. (2018) e reproduzi-lo o mais fielmente possível.
- Elaborar programas para a automatização da extração daquelas características de janelas de EEG.
- Automatizar o processo de preparação de experimentos usando as duas formas propostas de escolha de janelas para treino do sistema.
- Dos resultados de acurácia dos experimentos realizados, avaliar a influência sobre eles da forma de escolha das janelas para treino.

ATIVIDADES REALIZADAS

Atividade 1 – Estudo bibliográfico: Estudo da forma de separação de sequências de vetores utilizada por Tsiouris et al. (2018). Definição do chamado método de Tsiouris e método recomendado para separação de sequências.

Atividade 2 - Criação de sequências: Criação e utilização de programas para criação de sequências de vetores com as mesmas características utilizadas por Tsiouris et al. (2008) variando o número de vetores por sequência.

Atividade 3 - Testes iniciais: Treino e teste com as janelas criadas utilizando o método de separação de janelas tanto de Tsiouris et al. (2008) quanto o método criado pelo nosso grupo e medição de estatísticas de testes.

RESULTADOS OBTIDOS

Os dados utilizados foram obtidos na base de dados aberta CHB MIT Scalp EEG Database disponível no site PhysioNet (PHYSIONET, 2010) inicialmente descrito por Shoeb (2009). Esses exames foram baixados no formato European Data Format (EDF) e sofreram diversas etapas de pré-processamento antes de serem

utilizados na rede neural.

A primeira etapa foi a de conversão desses arquivos, utilizando o software livre EDFBrowser transformamos os arquivos do formato EDF para o formato *Comma separated values* (CSV). Em seguida, utilizando programas já em mãos do grupo de pesquisa, criados em outros trabalhos, extraímos as janelas interictais e pré-ictais dos exames e dessas janelas extraímos as características relatadas por Tsiouris et. al.: características de correlação, de Teoria de Grafos, de Transformada de Wavelet, de Transformada de Fourier e momentos estatísticos. Por fim, com tais vetores de características montados, foram montadas sequências contendo 15, 30, 45 ou 60 desses vetores finalizando assim o pré processamento.

Para cada paciente foram realizados dois experimentos para as janelas com vetores de Teoria de Grafos e para janelas com vetores de correlação, totalizando então 4 experimentos por paciente. O primeiro experimento foi treinar uma rede neural do tipo LSTM utilizando a forma de separação do conjunto de treino e de teste utilizada por Tsiouris et. al.. Já o segundo experimento foi o treinamento de uma rede neural, com mesma topologia, porém utilizando o método de separação recomendado pelo grupo. Esse novo método remove a mistura de janelas, fazendo assim com que seja impossível que duas janelas de uma mesma crise estejam uma no conjunto de treino e outra no conjunto de teste, a figura 1 representa a separação pelo método de Tsiouris et. al. e a figura 2 representa o método recomendado.

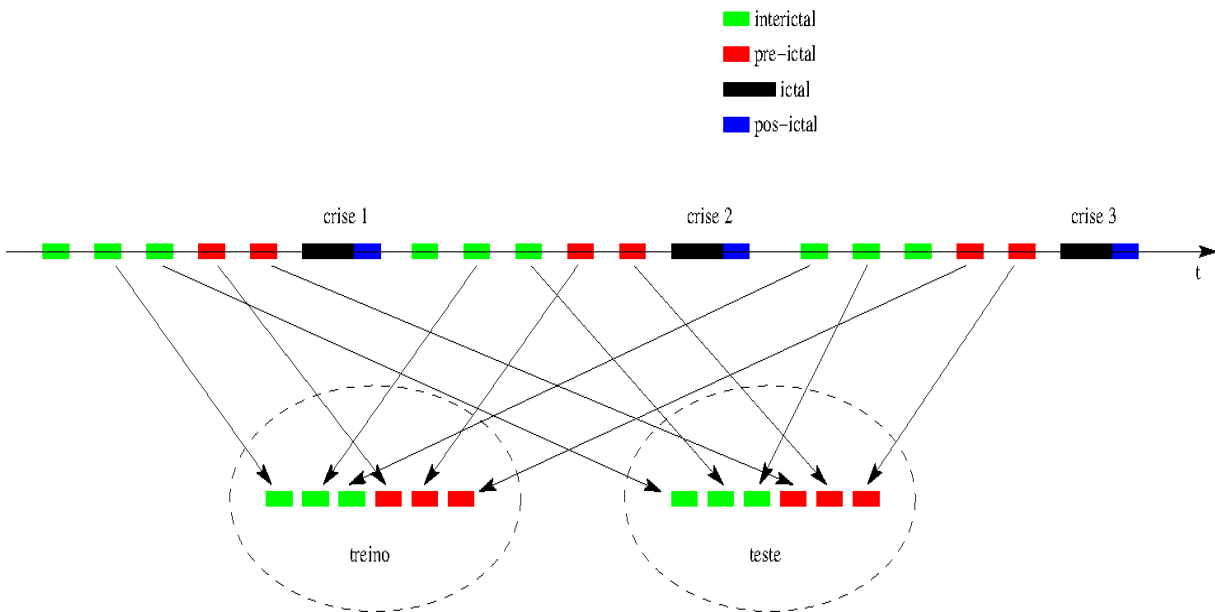


Figura 1: Separação de vetores segundo método de Tsiouris et. al. (Fonte: Professor Ricardo Pires)

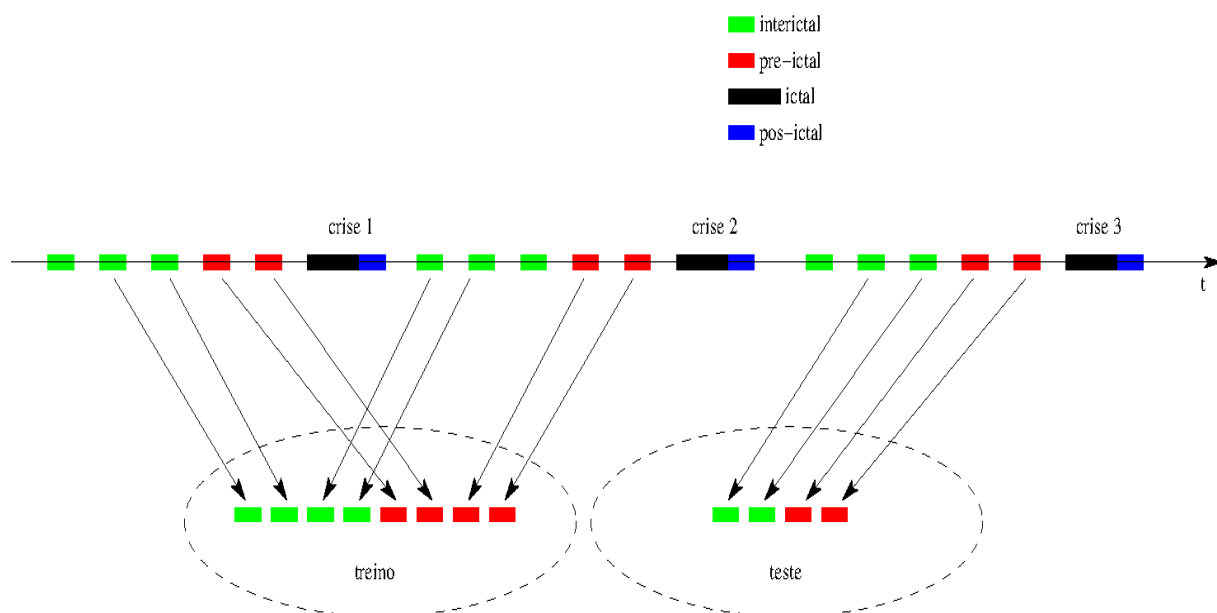


Figura 2: Separação de janelas segundo o método Recomendado (Fonte: Professor Ricardo Pires)

Os resultados desses testes vão de acordo com a hipótese proposta: como algumas sequências de vetores das crises do conjunto de teste estão no conjunto de treino alguma informação que a rede não poderia obter no treino é obtida, por isso os resultados nos experimentos pelo método de Tsiouris et. al. são muito melhores que os resultados do método proposto. A tabela 1 é uma comparação dos métodos de separação de sequências para alguns pacientes.

Paciente - quantidade de vetores	Método de separação	Tipos de vetores	Acurácia	Sensibilidade	Especificidade
CHB01 - 15 vetores	Método de Tsiouris	Correlação	100,00%	100,00%	100,00%
		Teoria de Grafos	53,68%	100,00%	0,00%
	Método recomendado	Correlação	62,11%	55,32%	68,75%
		Teoria de Grafos	50,53%	0,00%	100,00%
CHB07 - 45 vetores	Método de Tsiouris	Correlação	100,00%	100,00%	100,00%
		Teoria de Grafos	96,43%	92,31%	100,00%
	Método recomendado	Correlação	63,33%	40,00%	86,67%
		Teoria de Grafos	60,33%	33,33%	86,67%

Tabela 1: Comparação de resultados de testes (Fonte: Autor)

Vale notar que por se tratar de um problema de classificação com duas classes (pré ou interictal) a pior acurácia possível seria de 50% quando o algoritmo simplesmente decidisse aleatoriamente todas as sequências como pertencente a uma das duas classes, assim a acurácia de 60% não é boa. O maior indicativo de que isso ocorreu é quando a sensibilidade é nula e a especificidade é 100% ou vice versa, indicando que a rede ou classificou todos como positivos (pré-ictal) ou todos como negativos (interictal).

CRONOGRAMA FINAL

Os resultados parciais obtidos validam a hipótese do nosso grupo de pesquisa, assim os próximos passos se baseiam em realizar esses mesmos experimentos porém com os vetores com os outros tipos de características, além da preparação do artigo, a tabela 2 mostra o cronograma final para o projeto.

Julho - Agosto	Experimentos com conjuntos de vetores
Setembro - Outubro	Escrita do artigo
Novembro	Escrita do relatório final

Tabela 2: Cronograma da etapa final do projeto (Fonte: Autor)

REFERÊNCIAS

EPILEPSY FOUNDATION. **Epilepsy Foundation**. Disponível em: <<https://www.epilepsy.com/>>. Acesso em: 30 maio. 2021.

GADHOUMI, K. et al. Seizure prediction for therapeutic devices: A review. **Journal of Neuroscience Methods**, v. 260, p. 270–282, fev. 2016.

HOCHREITER, S.; SCHMIDHUBER, J. Long Short-Term Memory. **Neural Computation**, v. 9, n. 8, p. 1735–1780, 1 nov. 1997.

MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. **EEG (electroencephalogram)**. Disponível em: <<https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>>. Acesso em: 30 maio. 2021.

MULA, M.; MONACO, F. Ictal and Peri-Ictal Psychopathology. **Behavioural Neurology**, v. 24, n. 1, p. 21–25, 1 jan. 2011.

PHYSIONET. **CHB-MIT Scalp EEG Database**. Disponível em: <<https://physionet.org/content/chbmit/1.0.0/>>.

SHAFER, P., O.; SIRVEN, J., I. **About Epilepsy: The Basics**. Disponível em: <<https://www.epilepsy.com/learn/about-epilepsy-basics>>. Acesso em: 30 maio. 2021.

SHOEB, A. H. **Application of machine learning to epileptic seizure onset detection and treatment**. Tese (Doutorado)—[s.l.] Massachusetts Institute of Technology, 2009.

TSIOURIS, K. M. et al. A Long Short-Term Memory deep learning network for the prediction of epileptic seizures using EEG signals. **Computers in Biology and Medicine**, v. 99, p. 24–37, ago. 2018.

WORLD HEALTH ORGANIZATION. **Epilepsy**, 2019. Disponível em:
<<https://www.who.int/news-room/fact-sheets/detail/epilepsy>>. Acesso em: 30 maio. 2021

Assinatura do orientador: _____

Assinatura do bolsista: _____