



III LATIN AMERICAN WORKSHOP ON COMPUTATIONAL NEUROSCIENCE

Paulo Rogério de Almeida Ribeiro

Vinícius Rosa Cota

Alex Oliveira Barradas Filho

Dante Augusto Couto Barone

**Proceedings of the III Latin American Workshop on
Computational Neuroscience (LAWCN 2021)**



UNIVERSIDADE FEDERAL DO MARANHÃO



Prof. Dr. Natalino Salgado Filho
Reitor
Prof. Dr. Marcos Fábio Belo Matos
Vice-Reitor

EDITORA DA UNIVERSIDADE FEDERAL DO MARANHÃO



Prof. Dr. Sanatiel de Jesus Pereira
Diretor

CONSELHO EDITORIAL

Prof. Dr. Luís Henrique Serra
Prof. Dr. Elídio Armando Exposto Guarçoni
Prof. Dr. André da Silva Freires
Prof. Dr. Jadir Machado Lessa
Profª. Dra. Diana Rocha da Silva
Profª. Dra. Gisélia Brito dos Santos
Prof. Dr. Marcus Túlio Borowiski Lavarda
Prof. Dr. Marcos Nicolau Santos da Silva
Prof. Dr. Márcio James Soares Guimarães
Profª. Dra. Rosane Cláudia Rodrigues
Prof. Dr. João Batista Garcia
Prof. Dr. Flávio Luiz de Castro Freitas
Bibliotecária Suênia Oliveira Mendes
Prof. Dr. José Ribamar Ferreira Junior



Associação Brasileira
das Editoras Universitárias

Paulo Rogério de Almeida Ribeiro

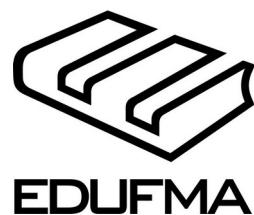
Vinícius Rosa Cota

Alex Oliveira Barradas Filho

Dante Augusto Couto Barone

**Proceedings of the III Latin American Workshop on
Computational Neuroscience (LAWCN 2021)**

São Luís



2022

Revisão

Prof. Dr. Paulo Rogério de Almeida Ribeiro (UFMA)
Prof. Dr. Vinícius Rosa Cota (Universidade Federal de São João del-Rei, UFSJ)
Prof. Dr. Alex Oliveira Barradas Filho (UFMA)
Prof. Dr. Dante Augusto Couto Barone (Universidade Federal do Rio Grande do Sul, UFRGS)

Projeto Gráfico

Prof. Dr. Paulo Rogério de Almeida Ribeiro (UFMA)
Prof. Dr. Vinícius Rosa Cota (Universidade Federal de São João del-Rei, UFSJ)

Dados Internacionais de Catalogação na Publicação (CIP)

Latin American Workshop on Computational Neuroscience (LAWCN 2021) (3.: 2021: São Luís, MA).

Proceedings of the III Latin American Workshop on Computational Neuroscience / Organization: Paulo Rogério de Almeida Ribeiro... [et al.]. — São Luís: EDUFMA, 2022.

91 p.

ISBN: 978-65-5363-087-1

1. Neurociências computacional. 2. Inteligência artificial. 3. Neuroengenharia. I. Ribeiro, Paulo Rogério Almeida. II. Cota, Vínicius Rosa. III. Barradas Filho, Alex Oliveira. IV. Barone, Dante Augusto Couto. V. Título.

CDD 006.326.168
CDU 004.032.26:616.8

Bibliotecária: Regina França Cutrim CRB 13 / 542

EDUFMA | Editora da UFMA
Av. dos Portugueses, 1966 – Vila Bacanga
CEP: 65080-805 | São Luís | MA | Brasil
Telefone: (98) 3272-8157
www.edufma.ufma.br | edufma@ufma.br

Overview

The Latin American Workshop on Computational Neuroscience (LAWCN) was born in 2017 and it is devoted to address themes such as Computational Neuroscience, Artificial Intelligence, Neuroscience, and Neuroengineering. It has been biennially holding since: I LAWCN (2017) in Porto Alegre (Brazil) and II LAWCN (2019) in São João Del-Rei (Brazil). Due to the COVID-19 pandemic, its third edition, III LAWCN'21, was held as a hybrid event, i.e. online and in person in the city of São Luís (Maranhão, Brazil) during 8th to 10th December 2021.

This proceedings contains some works (oral presentation and posters) that were presented at the LAWCN'21. It has works on Brain-Computer Interface, Machine learning to predict epilepsy seizures, Machine learning for electrocardiogram analysis, Model for Prefrontal Activity, Alzheimer disease, Parkinson's Disease, Conductance based integrate and fire neurons, Izhikevich model simulation and so on.

LAWCN'21 was organized by Federal University of Maranhão (UFMA – *Universidade Federal do Maranhão*, Brazil) with support of *Fundação Sousândrade*. (FSADU). Financial aid was provided Plexon Inc and by the International Brain Research Organization (IBRO). NeuroTechX (NTX) was our partner and contributed with ideas, minds, panelists, and worldwide dissemination.

We, the editors of this proceedings would like to thank: UFMA, FSDAU, Plexon Inc, IBRO, NTX, all keynote speakers, authors, PC members (reviewers) and participants.

Paulo Rogério de Almeida Ribeiro
Vinícius Rosa Cota
Alex Oliveira Barradas Filho
Dante Augusto Couto Barone

Organization



Sponsors and Support



EDITOR

Paulo Rogério de Almeida Ribeiro
Federal University of Maranhão (UFMA) – Brazil

CO-EDITORS

Vinícius Rosa Cota
Federal University of São João del-Rei (UFSJ) – Brazil

Alex Oliveira Barradas Filho
Federal University of Maranhão (UFMA) – Brazil

Dante Augusto Couto Barone
Federal University of Rio Grande do Sul (UFRGS) – Brazil

ORGANIZED BY

Federal University of Maranhão (UFMA) - Brazil

CONFERENCE CHAIRS

Paulo Rogério de Almeida Ribeiro (Chair)
Federal University of Maranhão (UFMA) – Brazil

Vinícius Rosa Cota (Co-chair)
Federal University of São João del-Rei (UFSJ) – Brazil

ORGANIZING COMMITTEE

Alex Oliveira Barradas Filho
Federal University of Maranhão (UFMA) – Brazil

Alexandre César Muniz de Oliveira
Federal University of Maranhão (UFMA) – Brazil

Allan Kardec Duailibe Barros Filho
Federal University of Maranhão (UFMA) – Brazil

Dante Augusto Couto Barone
Federal University of Rio Grande do Sul (UFRGS) – Brazil

Fabrício Lima Brasil
Edmond and Lily Safra International Institute of Neuroscience, Santos Dumont Institute (ISD) – Brazil

Frederico Augusto Casarsa de Azevedo
Massachusetts Institute of Technology (MIT) & Harvard Medical School – USA

Jaime Andres Riascos Salas
Institución Universitaria de Envigado – Colombia

Paulo Rogério de Almeida Ribeiro
Federal University of Maranhão (UFMA) – Brazil

Sen Cheng
Ruhr University Bochum – Germany

Sofia Maria Amorim Falco Rodrigues
Federal Center for Technological Education of Minas Gerais (CEFET-MG) – Brazil

Vinícius Rosa Cota
Federal University of São João del-Rei (UFSJ) – Brazil

PROGRAM COMMITTEE CHAIR

Paulo Rogério de Almeida Ribeiro
Federal University of Maranhão (UFMA) – Brazil

PROGRAM COMMITTEE

Abner Cardoso Rodrigues Neto
Edmond and Lily Safra International Institute of Neuroscience, Santos Dumont Institute (ISD) –
Brazil

Alan Talevi
National University of La Plata – Argentina

Alex Oliveira Barradas Filho
Federal University of Maranhão (UFMA) – Brazil

Alexandre César Muniz de Oliveira
Federal University of Maranhão (UFMA) – Brazil

Ali Khaledi Nasab
Stanford University – USA

Allan Kardec Duailibe Barros Filho
Federal University of Maranhão (UFMA) – Brazil

Ander Ramos-Murguialday
University of Tübingen – Germany

André Salles Cunha Peres
Universidade de Coimbra – Portugal

Andrea Sarasola
TECNALIA - Neurotechnology Laboratory – Spain

Andreea-Ioana Sburlea
Graz University of Technology – Austria

Antônio Roque
University of São Paulo (USP) – Brazil

Bruno Feres de Souza
Federal University of Maranhão (UFMA) – Brazil

Carlos Campos
University of Itaúna – Brazil

Carlos Dias Maciel
University of São Paulo (USP) – Brazil

Carolina Bellera
National University of La Plata – Argentina

Cesar Alberto Collazos
Universidad del Cauca – Colombia

Cristina Simon-Martinez
University of Applied Sciences Western Switzerland – Switzerland

Dania Gutiérrez
Cinvestav Monterrey – Mexico

Dante Augusto Couto Barone
Federal University of Rio Grande do Sul (UFRGS) – Brazil

Dimitris Kugiumtzis
Aristotle University of Thessaloniki – Greece

Dinesh Kant Kumar
Royal Melbourne Institute of Technology (RMIT) University – Australia

Edgard Morya
Edmond and Lily Safra International Institute of Neuroscience, Santos Dumont Institute (ISD) – Brazil

Edison Pignaton de Freitas
Federal University of Rio Grande do Sul (UFRGS) – Brazil

Eduardo López-Larraz
Bitbrain – Spain

Erivelton Geraldo Nepomuceno
Federal University of São João del-Rei (UFSJ) – Brazil

Fabrício Lima Brasil
Edmond and Lily Safra International Institute of Neuroscience, Santos Dumont Institute (ISD) – Brazil

Frederico Augusto Casarsa de Azevedo
Massachusetts Institute of Technology (MIT) & Harvard Medical School – USA

Gabriel Mindlin
Universidad de Buenos Aires – Argentina

Gabriela Castellano
University of Campinas (Unicamp) – Brazil

Gil Lopes
University Institute of Maia (ISMAI) – Portugal

Guillermo Cecchi
IBM T.J. Watson Research Center – USA

Héctor Julián Tejada Herrera
Federal University of Sergipe (UFS) – Brazil

Jaime Andres Riascos Salas
Institución Universitaria de Envigado – Colombia

Jean Faber
Federal University of São Paulo (UNIFESP) – Brazil

Jessica Cantillo-Negrete
National Institute of Rehabilitation – Mexico

Jim Tørresen
University of Oslo – Norway

Jorge Parraga-Alava
Universidad Técnica de Manabi – Ecuador

José C. Príncipe
University of Florida – USA

Jose Donoso
Ruhr University Bochum – Germany

Juan Felipe Ramirez Villegas
Institute of Science and Technology Austria – Austria

Justus Alfred Kromer
Stanford University – USA

Kai Olav Ellefsen
University of Oslo – Norway

Leonardo Bonato Félix
Federal University of Viçosa (UFV) – Brazil

Lilian Konicar
Medical University of Vienna – Austria

Lucas Alberca
National University of La Plata – Argentina

Luciana Gavernet
National University of La Plata – Argentina

Marcelo Pias
Federal University of do Rio Grande (FURG) – Brazil

Marianna Semprini
Istituto Italiano di Tecnologia – Italy

Michela Chiappalone
Istituto Italiano di Tecnologia – Italy

Nerea Irastorza
TECNALIA - Neurotechnology Laboratory – Spain

Nivaldo Antônio Portela de Vasconcelos
Federal University of Pernambuco (UFPE) – Brazil

Omar Andrés Carmona Cortés
Federal Institute of Maranhão (IFMA) – Brazil

Patricio Orio
Universidad de Valparaíso – Chile

Paulo Rogério de Almeida Ribeiro
Federal University of Maranhão (UFMA) – Brazil

Pedro Almeida
University of Lisboa – Portugal

Reinhold Scherer
University of Essex – UK

Renan Cipriano Moioli
Federal University of Rio Grande do Norte (UFRN) – Brazil

Rodrigo Alejandro Sierra Ordoñez
University of Szeged – Hungary

Sen Cheng
Ruhr University Bochum – Germany

Ulysse Côté-Allard
University of Oslo – Norway

Vinícius Rosa Cota
Federal University of São João del-Rei (UFSJ) – Brazil

KEYNOTE SPEAKERS

Alan Talevi
National University of La Plata – Argentina

André Carvalho
University of São Paulo – Brazil

Angela Wyse
Federal University of Rio Grande do Sul – Brazil

José C. Príncipe
University of Florida – USA

Kerstin Schmidt
Federal University of Rio Grande do Norte – Brazil

Peter Tass
Stanford University – USA

Contents

Abstracts

| | |
|--|--------|
| Lower Extremity Rehabilitation of Chronic Stroke Patients using a Brain-Computer Interface System <i>Marc Sebastian-Romagosa; Woosang Cho, Rupert Ortner; Alexander Lechner; Christoph Guger</i> | ... 1 |
| Machine learning models in the study and prediction of epilepsy <i>Gabriel Giovanaz; Marco Idiart</i> | ... 2 |
| Automatic classifier for pattern recognition in epilepsy in electroencephalographic recordings <i>Pedro H.A.B. Santos; Jasiara C. Oliveira; Vinícius R. Cota; Sofia M.A.F. Rodrigues</i> | ... 3 |
| Classificação automática de distúrbios elétricos <i>Thalita K. Pereira; Márcio W. Santana; Sofia M. A. F. Rodrigues</i> | ... 4 |
| A Model for Prefrontal Activity During Moral Decision Making in Human and Non-Human Primates <i>César Daniel Alves Caldeira; Vinícius Rezende Carvalho; Renato César Cardoso</i> | ... 5 |
| Motor Adaptation to Unpredictable Contexts are Influenced by the Ability Level and the Type of Training <i>Cíntia O. Matos; Carlos E. Campos; Crislaine R. Couto; Lucas C. C. Silva; Paulo R. A. Ribeiro; Herbert Ugrinowitsch</i> | ... 6 |
| Study of structural and functional differences in the brain of healthy elderly bilinguals and those with Alzheimer's disease <i>Mariana de Melo G. Aguiar; Marco Antonio G. Carvalho</i> | ... 7 |
| Non-periodic stimulation of the subthalamic nuclei as an alternative treatment for Parkinson's Disease <i>Isabela Colem Castelo Borges; Matheus Cavalcante Meira; Luan Alves Pereira; Daniel de Castro Medeiros; Márcio Flávio Dutra Moraes; Laila Laila Cristina Moreira Damázio; Vinícius Rosa Cota</i> | ... 8 |
| Um modelo computacional da circuitaria hipocampal baseado no Neuronify(R) para ensino de neurociência <i>Lucas Ferreira; Jean Faber</i> | ... 9 |
| A mean-field model for conductance based integrate and fire neurons with variable timescales <i>Marcelo P. Becker; Marco A. P. Idiart</i> | ... 10 |

Spatial Allocation of Memories in Modules using Theta-Gamma Traveling Waves ... 11

Gustavo D. Soroka; Marco Idiart

Izhikevich model simulation of spiking neurons in a Web-based architecture ... 12

Rodrigo de Sousa Gomide; Marcus Fraga Vieira

Microscale modeling of seizures by spatiotemporal chaos in coupled map lattices ... 13

Rafael V. Stenzinger; Marcelo H. R. Tragtenberg

Avaliação topológica sobre padrões espaço-temporais de mortalidade por COVID-19 em função de fatores socioeconômicos no estado de São Paulo ... 14

Rodrigo Lantyer; Priscila Antoneli; Marimélia Porcionatto; Jean Faber

Papers

Machine learning-based application for long-term electrocardiogram analysis ... 15

Juliana Mycaelle Silva; Jonathan Araujo Queiroz; Luís Filipe da Silva; Gean C. Sousa; Marta de Oliveira; Allan Kardec Barros

Modelo baseado em aprendizado de máquinas para classificar atividades encefálicas cognitivas ... 33

Juliana Mycaelle Silva; Gean C. Sousa, George V. S. Souza; Jonathan A. Queiroz; Allan Kardec Barros

A Random LSTM Model for Stock Market Prediction ... 48

Elaine Pinto Portela; Omar Andres Carmona Cortes

COVID-19 Cognitive Sequelae and Their Possible Relation with Educational Issues: a systematic review ... 62

Drielle Viana Vieira; Tainá dos Santos Rêgo; Sayonara Pereira da Silva; Maria Carolina Gonzalez; Edgard Morya

Lower Extremity Rehabilitation of Chronic Stroke Patients using a Brain-Computer Interface System

Marc Sebastian-Romagosa¹; Woosang Cho¹, Rupert Ortner¹; Alexander Lechner²; Christoph Guger^{1,2}

¹g.tec medical engineering Spain S.L., Barcelona, Spain

²g.tec medical engineering GmbH, Schiedlberg, Austria

Introduction: Neurorehabilitation based on Brain-Computer Interfaces (BCIs) show important rehabilitation effects for patients after stroke. Previous studies have shown improvements for patients that are in a chronic stage and/or have severe hemiparesis and are particularly challenging for conventional rehabilitation techniques. The effectiveness of this BCI system for upper limb rehabilitation was proven in a clinical study with a sham group. The goal of this study is to prove the concept for the lower extremities too.

Methods: For this publication ten stroke patients in chronic phase with hemiparesis in the lower extremity were recruited. All of them participated in 25 BCI sessions about 3 times a week. The BCI system (see Fig 1.) was based on the Motor Imagery (MI) of the paretic ankle dorsiflexion and healthy wrist dorsiflexion with Functional Electrical Stimulation (FES) and avatar feedback. Assessments were conducted to assess the changes in motor improvement before, after and during the rehabilitation training. Our primary measures used for the assessment were 10-meters walking test (10MWT), Range of Motion (ROM) of the ankle dorsiflexion and Timed Up and Go (TUG).

Results: Results show a significant increase in the gait speed in the primary measure 10MWT fast velocity of 0.16 m/s (SD = 0.14). This improvement is above of the minimally clinically important difference (MCID). The speed in the TUG was also significantly increased by 0.06 m/s, P = 0.002. One patient was not able to perform TUG assessment before the rehabilitation training but was able to perform it after the BCI treatment with time 92.2 seconds. The passive ROM assessment increased 8.61° (SD = 6.54), P = 0.002, and active ROM increased 8.50° (SD = 7.23) after rehabilitation training, P = .008.

Discussion: These outcomes show the feasibility of this BCI approach for chronic stroke patients, and further support the growing consensus that these types of tools might develop into a new paradigm for rehabilitation tool for stroke patients. However, the results are from only ten chronic stroke patients so the authors believe that this approach should be further validated in broader randomized controlled studies involving more patients. MI and FES-based non-invasive BCIs are showing improvement for the gait rehabilitation of the patients in the chronic stage after stroke. This could have an impact on the rehabilitation techniques used for these patients.

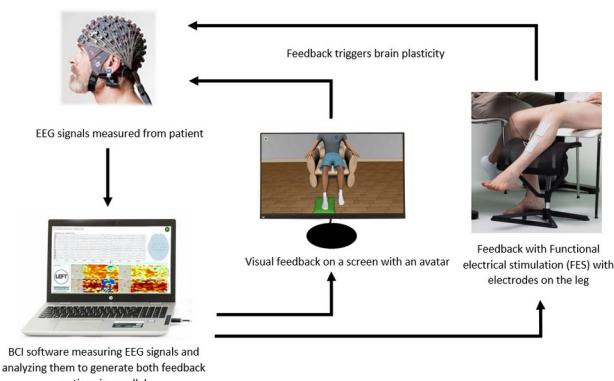


Fig. 1. BCI system

Machine learning models in the study and prediction of epilepsy seizures

Gabriel Giovanaz¹; Marco Idiart¹

¹Universidade Federal do Rio Grande do Sul, Brazil

Introduction: Epilepsy is a disease that affects approximately 50 million of people worldwide; it's characterized by spontaneous recurrent seizures that can go undetected of different types that can impair the quality of life of these patients. It is estimated that about one third of patients with epilepsy have refractory seizures: seizures resistant to available medication. The refractory epilepsy can cause damage to the brain, affecting their quality of life. For some of these patients, the surgical treatment this affected brain tissue removal can help to decrease de seizure frequency. For the patients with refractory epilepsy, prediction of seizures would be important to find a way to stop the occurrence of the seizures.

Methods: Seizure prediction techniques using mathematical modeling and machine learning have been explored with good results, even though not always with an online model that can predict these seizures in real time. In this work we assess the performance of a deep learning model trained with EEG data from CHB-MIT, a dataset with 24 patients, with seizures onset already identified by the authors of the data. This work utilizes different preprocessing techniques to predict the occurrence of these seizures. The models were trained so that the prediction could be made within a 30-minute interval before its occurrence.

Results: 4 different models were trained with different preprocessing techniques. Average AUC score of the models reached 81.5%, 80.5%, 80.7% and 94.3%, reaching 100% AUC score in some subjects for the last model.

Discussion: The results show that it is possible to predict these seizures with both precision and in real time, enabling the engineering of gadgets that can be carried by people and help them in the identification and management of seizures.

Acknowledgements: This project would not have been possible without fundings from CAPES, whose resources fund the UFRGS Neuroscience graduate program.

Automatic classifier for pattern recognition in epilepsy electroencephalographic recordings

Pedro H.A.B. Santos¹; Jasiara C. Oliveira²; Vinícius R. Cota³; Sofia M.A.F. Rodrigues¹

¹Centro Federal de Educação Tecnológica de Minas Gerais, Nepomuceno, Brasil

²Centro Universitário Presidente Tancredo de Almeida Neves, São João del-Rei, Brasil

³Universidade Federal de São João del-Rei, São João del-Rei, Brasil

Introduction: Epilepsy is a public health issue worldwide with a significant portion of non-treated patients and our group has shown that a non-periodic stimulation form applied to the basolateral amygdala is able to suppress seizures in epilepsy animal models [1]. In addition closed-loop systems that provide patient monitoring had been carried out over last years including or not electrical stimulation to anticipate a seizure, using simple artificial neural networks or more sophisticated machine learning systems [2]. One way to measure neural synchronism for feature extraction can be through the interaction of neural oscillations in different frequency bands [3] or assessing distinguishable electrographic patterns, such as epileptiform spikes [4]. In this sense, the aim of this study is obtaining a possible predictor classifier to healthy and seizure states using a Multilayer Perceptron (MLP) with optimized training by a Genetic Algorithm (GA).

Methods: Dataset consists of local field potentials (LFP) recorded from cortex, hippocampus, and thalamus (2000 V/V gain, 1000 Hz sampling rate, and band-pass filtered 0.3 to 300 Hz) from a total of 42 male Wistar rats divided into control and NPS stimulated groups (JCO PhD experiment). Controlled intravenous infusion of PTZ provided a gradual transition of susceptibility states, until onset of forelimb myoclonus (MYO - chosen ictal state in this work). The chosen preictal state consists of 30 seconds before MYO onset. Modulation Index (MI) (assessing 0 to 14 Hz for phase and 8 to 300 Hz for amplitude combinations to measure phase-amplitude coupling), epileptiform spike count, and spike temporal coincidence count were used to feature extraction (see ref 3 and 4) for MLP input. Spike coincidence was assessed by counting the number of temporal coincidences between detected epileptiform spikes in a pair of LFP channels, in this work time window fixed in 0.001 to 0.05 s. Both network and GA were developed with PyGAD library, in Python, doing 30 GA executions to obtain the best network adjustment for test phase and statistical validation. GA setting: randomized initialization, n = 100 (pop size); roulette wheel parent selection; scattered crossover; randomized mutation (10% probability rate); only offspring to next generation and 5000 generations maximum.

Results: The best general accuracy in a MLP test was obtained with only MI feature extraction (87.736%). On the other hand best results in terms of healthy, preictal, and ictal state recognition were obtained using all chosen feature extraction tools. Low false positive rates (0% - median - healthy versus ictal and 4.545% - median - healthy versus preictal and preictal versus ictal) and reasonable accuracy (83% - median) considering control versus NPS group recognition (for all evaluated states).

Discussion: Classifiers based on computationally simple neural networks with Wavelet transform feature extraction have been showing potential results [2]. Thus, it is clear the better efficiency of an automatic classifier for possible seizure anticipation can be linked to effective feature extraction for assessing the expected synchronism and desynchronism moments part of important states in epilepsy context (preictal and ictal phases).

Acknowledgements: IBRO (LAWCN Financial Support). VRC (co-advisor) who designed the experimental protocol and managed fundings for animal experimentation.

References: [1] doi 10.1016/j.yebeh.2019.106609 [2] doi 10.1016/j.trit.2016.08.001 [3] doi 10.1152/jn.00106.2010 [4] doi 10.1007/978-3-030-36636-0 11

Classificação automática de distúrbios elétricos

Thalita K. Pereira¹; Márcio W. Santana¹; Sofia M. A. F. Rodrigues¹

¹Centro Federal de Educação Tecnológica de Minas Gerais, Nepomuceno, Brasil

Introdução: Com a integração de sistemas de energia, há cada vez mais riscos à qualidade da energia elétrica (QEE) em diferentes estágios, de produção, transformação, entrega e consumo desta. Assim, a classificação automática de possíveis distúrbios que podem surgir em parâmetros do sistema elétrico de potência pode ser a base para lidar com os problemas da QEE [1]. Nesse sentido, o uso de tecnologias como as redes neurais artificiais pode ser uma saída para o reconhecimento automático desses distúrbios [1,2], aliando ainda a um processo de treinamento otimizado, via Algoritmos Genéticos (AG) [3]. Diante disto, o objetivo deste trabalho é desenvolver um classificador automático em Python para alguns dos possíveis distúrbios que impactam na QEE, a partir de uma rede Perceptron Multicamadas (MLP – *Multilayer Perceptron*), com processo de treinamento otimizado, via AG.

Métodos: Os dados utilizados foram obtidos computacionalmente via MATLAB®, com frequência fundamental de 60 Hz e amostragem de 15360 Hz, em 256 amostras por ciclo e relação sinal-ruído de 60 dB, um total de 10000 amostras divididas igualmente em 5 classes: cada distúrbio analisado da QEE analisado (afundamento da tensão, transitórios oscilatórios, interrupção curta e harmônicos) e o sinal sem distúrbios. A extração de características para entrada da RNA foi feita com o filtro Hodrick-Prescott, a Transformada Rápida de Fourier e pela raiz do valor quadrático médio (valor eficaz do sinal). Na implementação do algoritmo proposto utilizou-se a biblioteca PyGAD, em Python, tanto para a RNA quanto para o AG, sendo a RNA formada por 6 neurônios na camada de entrada, 15 neurônios internos e 5 neurônios na de saída. No AG representou-se os indivíduos de forma binária, sendo a população inicial gerada aleatoriamente, com o cruzamento ocorrendo através de pais escolhidos via roleta, probabilidade de mutação de 10%, a ocorrer com genes aleatórios, nova geração formada pelos descendentes e 5000 gerações como critério de parada, para cada uma das 5 execuções feitas do algoritmo. A melhor solução obtida pelo AG corresponde à melhor configuração da RNA, a ser utilizada para os testes finais.

Resultados: Considerando os resultados do teste da MLP onde foi obtida a melhor acurácia geral, 86,92%, na comparação do reconhecimento entre os distúrbios analisados e o sinal normal, observou-se acurárias de quase 100%, com exceção do afundamento de tensão (68,025%). Ainda neste teste, na classificação entre os possíveis distúrbios analisados, as maiores dificuldades foram na separação entre um destes e o afundamento de tensão ($\approx 67\%$), embora, em contrapartida, foram obtidas novamente acurárias de quase 100% para o reconhecimento entre os demais distúrbios.

Discussão: Utilizando redes neurais convolucionais, por exemplo, em modelos de pouco custo computacional existem evidências recentes de acurárias gerais muito próximas a 100% [1], entretanto para efetiva comparação e validação com o algoritmo proposto é necessária a realização de mais execuções e o desenvolvimento de análise estatística. Modificações simples no algoritmo proposto como o uso de maior quantidade de neurônios na camada interna ou outros ajustes no AG como uma mudança na forma de seleção de soluções para o processo de geração de outras novas podem surtir em melhoria na detecção do afundamento de tensão e, consequentemente, no aumento da acurácia geral, embora caiba aqui analisar a efetividade da extração de características frente o reconhecimento do afundamento de tensão, principalmente.

Referências: [1] doi 10.1016/j.apenergy.2018.09.160 [2] doi 10.1016/j.epsr.2006.12.011 [3] doi 10.1016/S0305-0483(99)00027-4

A Model for Prefrontal Activity During Moral Decision Making in Human and Non-Human Primates

César Daniel Alves Caldeira¹; Vinícius Rezende Carvalho¹; Renato César Cardoso¹

¹Universidade Federal de Minas Gerais, BRA

Introduction: Moral decision-making in the brain emerges from two main structures. The ventromedial prefrontal cortex (VMPFC) and dorsolateral (DLPFC) cortex. The VMPFC is associated with the regulation of emotional processes with greater activation related to deontological decisions. The DLPFC activity, on the other hand, is linked to cognitive control, resulting in consequentialist behavior. Chimpanzees (*Pan troglodytes*) and bonobos (*Pan paniscus*) are our closest living relatives. Although belonging to the same genus, they have behaviors that are very different. Recording brain activity in these animals poses great challenges, making computational models of neuronal activity invaluable tools to study these species. That is, models can be developed that enable the manipulation of simulation parameters that represent the neuronal activity of human and non-human primates, under conditions that were previously unfeasible *in vivo*. The objective of this work was to evaluate differences in neuronal activity regimes of three neural networks under strong and weak stimuli. Each network is a simple representation of prefrontal activity in humans, chimpanzees, or bonobos. Stimuli are aimed to represent pro-consequentialist and pro-deontological stimuli, respectively.

Methods: The simulations were implemented in NEST (Neural Simulation Tool) and parameters were based on literature data regarding size (adjusted for the number of neurons used: human-VM:3400, DL:3200/bonobo-VM:840, DL:790/ chimpanzee-VM:933, DL:878) and network connectivity with the amygdala (human-VM:0.5, DL:0.2/Pan-VM:0.35, DL:0.18). The configuration of neurons was based on Brunel's neocortical network. Electrical stimuli with Poisson distribution were applied to networks in order to emulate the activity triggered by an evidence of the environment in which a decision would be required. The stimuli weights ranged from 1 to 100 spikes/s and applied for 1s. DC level (average voltage), power spectral density (PSD) and signal stationarity were analyzed.

Results: In humans, the increased activity in DLPFC occurred under intensities 40, 50 and 75, while in VMPFC under 50, 75 and 100. In Pan, all stimulus intensities resulted in high neuronal activity. The basal PSD analysis showed that the 3 species have high power in the Beta frequency range and specifically in Pan, the range is also high in Gamma. For stationarity, human DLPFC under 50 and chimpanzee DLPFC under 1, 5, 10 and 40 showed a non-stationary trend ($p>0.0001$) using the ADF test.

Discussion: The human-inspired model shows different neuronal activity regimes when compared with the two Pan models, which resemble each other but do not have identical behavior. The human DLPFC model displayed stationarity under a more limited input intensity range, when compared with the chimpanzee model. The non-stationarity of human DLPFC under intensity 50 may be related to the phenomenon of moral tendency transition, as well as the DLPFC for chimpanzees. The increased neuronal activity in Pan for all levels may be related to the excitability of neural networks from the basal level. The results show that differences in neuronal connectivity and network size are likely to play a role in moral decision making. Thus, this work presents an initial step for simulating the phenomenology of moral decision making in humans and other primates.

Motor Adaptation to Unpredictable Contexts are Influenced by the Ability Level and the Type of Training

Cíntia O. Matos¹; Carlos E. Campos^{1,2}; Crislaine R. Couto^{1,3}; Lucas C. C. Silva⁴; Paulo R. A. Ribeiro⁴; Herbert Ugrinowitsch¹

¹Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

²Universidade de Itaúna, Itaúna, Brazil

³Instituto Metodista Izabela Hendrix, Belo Horizonte, Brazil

⁴Universidade Federal do Maranhão, São Luís, Brazil

Introduction: The neuromotor system has the ability to update the repertory of control strategy in function of the performance stabilization. Thus, the objectives of this study were to investigate the effects of manipulating different levels of performance stabilization (stabilization and specialization) as a function of the structuring of practice (constant and random practices) and to investigate how control is performed in the acquisition and adaptation to unpredictable motor perturbations.

Method: The task consisted of moving the physical effector in order to intercept a virtual target projected on a screen. Participants (n=56) were divided into 4 groups and submitted to a pre-exposure phase, and 24 hours later to a phase of exposure to perturbations, which different weights were added to the physical effector (Etic protocol nº 27875514.5.0000.5149). The kinematic variables as well as the electromyographic activity of the agonist and antagonist muscles were captured to investigate control strategies.

Results: The results showed that practice constant with stabilization level had lower performance than other groups during acquisition and adaptation phases. Moreover, reaching the level of performance stabilization from constant and random practices leads to the formation of control structures that use different strategies during the skill acquisition process.

Discussion: Achieving the stabilization level of performance from random practice and performance specialization from constant practice and random practice seem to lead to the formation of control structures that use similar strategies and reflect in better performance levels in the face of unpredictable motor perturbations. Thus, the findings of this study allow us to conclude that not only the level of stabilization of performance, but also the way in which the practice is structured, influence the formation of structures of motor control and, consequently, motor adaptation.

Acknowledgements: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brazil (CAPES) – Finance code 001 (Universidade Federal de Minas Gerais).

Study of structural and functional differences in the brain of healthy elderly bilinguals and those with Alzheimer's disease

Mariana de Melo G. Aguiar¹; Marco Antonio G. Carvalho¹

¹School of Technology, University of Campinas, Limeira/SP, Brazil

Introduction: Alzheimer's disease (AD) is the leading cause of dementia amongst elderly people in the world. Currently, AD has neither cure nor disease-modifying drugs available to halt its progression. Given the elderly population growth in the last decades - and the consequent rise in AD prevalence - being able to identify activities that prevent or delay the symptoms' onset can be valuable from both social and economic viewpoints. Some studies have found evidence that bilingualism might be one of these beneficial activities. Therefore, we aim to study the differences in the aging brain of healthy bilinguals and AD bilingual patients compared to their monolinguals counterparts via systematic review.

Methods: For our systematic review, we defined the keyword combination (*mri OR fmri OR imag* OR pet OR dti*) AND (*bilingual* OR multilingual**) AND (*alzheimer OR reserve*) to use as search query in 5 databases, namely: BVS, Web of Science, Scopus, PsycINFO and Pubmed. Exclusion criteria encompass reviews, proceedings, commentary and letters articles and non-peer reviewed works. For inclusion, the study should use imaging techniques to assess structural or functional differences between older bilinguals and monolinguals. We adopted the *Joanna Briggs Institute Appraisal Tools* [1] to evaluate the risk of bias and follow the *PRISMA* [2] statement to present the findings. We clarify that this is still an ongoing systematic review.

Results: 453 records were identified in five different databases. We removed 105 duplicated records before the screening and other 314 records after the first screening. With 34 articles remaining, we were able to retrieve 33 of them and start the full article reading and assessment, which is our current stage.

Discussion: Since we are reviewing the selected papers, we are still gathering the available data. Nevertheless, synthesize and evaluate the evidence of differences in older bilingual and monolingual brain may help our understanding of activities that can aid a healthier aging.

Acknowledgements: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior -- Brasil (CAPES) – Funding Number 001.

References:

- [1] Moola, S., et al. Joanna Briggs Institute Reviewer's Manual. The Joanna Briggs Institute 5 (2017).
- [2] Page, Matthew J., et al. BMJ 372 (2021).

Non-periodic stimulation of the subthalamic nuclei as an alternative treatment for Parkinson's Disease

Isabela Colem Castelo Borges¹; Matheus Cavalcante Meira¹; Luan Alves Pereira¹; Daniel de Castro Medeiros²; Márcio Flávio Dutra Moraes³; Laila Laila Cristina Moreira Damázio¹; Vinícius Rosa Cota¹.

¹Universidade Federal de São João del Rei (UFSJ), Minas Gerais, Brazil.

²Neuroscience Department, Karolinska Institutet, Sweden.

³Núcleo de Neurociências (NNC), Universidade Federal de Minas Gerais (UFMG), Brazil.

Introduction: Deep brain stimulation is a well-established treatment for refractory Parkinsonisms and is usually administered in high-frequencies (>100 hZ) at the subthalamic nucleus. However, a low-frequency/low-energy stimulation protocol would be preferable due to a combination of reasons, such as a lower charge density in the electrodes (reducing risks of heat production and electroporation), a reduced battery consumption (which increases the time needed for surgical replacement interventions) and a decreased probability of tissue adaptation (which causes long-term effect loss). Taking into consideration the principles of open science practice, this work aims to disclose the investigation, in a Parkinson's disease (PD) animal model, the acute effects of the bilateral application on the subthalamic nuclei of a non-periodic (NP) and low-frequency stimulation, the NPS-IH. This novel stimulation parameter is entitled in account of the inverse histogram (IH) pattern obtained when graphically represented and has achieved, on epilepsy animal models, significant effects on the suppression of hypersynchronous neuronal firing (Cota *et al.*, 2019).

Methods: Wistar rats will undergo three stages of procedures: 1) stereotaxic surgery for both the electrical lesion of the substantia nigra (to induce the parkinsonism) and implantation of the stimulation and local field potential (LFP) recording electrodes; 2) functional stage for evaluation of motor performance and; 3) final stage for LFP recording. The coordinates for the subthalamic nuclei stimulation (AP: -3.6, ML: 2.6, DV: - 8.4, relative to the brain surface) were selected in order to prioritize the posterolateral region, while the chosen sites for LFP recording were the primary motor cortex (+2.2 mm, ML: +2.8 mm, DV: -2.2 mm, from bregma) and the dorsal striatum (+ 1.6 mm, ML: +2.4 mm, DV: -5.0 mm, from bregma). For the electrical lesion of the substantia nigra, a 1 mA current will be administered for 10 seconds. NPS-IH will be applied bilaterally as charge-balanced biphasic current pulses with low average frequency (4 pulses/second), 150 to 400 μ A of amplitude and 100 μ s of duration (Cota *et al.*, 2019). For motor evaluation, open field, parallel bars and grid-walking tests will be performed.

Discussion: Evidence of synchronization disturbances in PD are becoming increasingly robust. The experimental success observed at the application of NPS-IH on epilepsy animal models (Cota *et al.*, 2019), a classic disease in terms of synchronization disturbances, brings up high chances of its utilization for the PD's scenario, in which we expect to observe both restoration on LFP levels and symptoms improvement.

Acknowledgements: Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

References: Cota, V. R. *et al.* Epilepsy & Behavior, 05 November, 2019. DOI: <https://doi.org/10.1016/j.yebeh.2019.106609>

Um modelo computacional da circuitaria hipocampal baseado no Neuronify(R) para ensino de neurociência

Lucas Ferreira¹ and Jean Faber¹

¹Universidade Federal de São Paulo

Introdução: A formação hipocampal é considerada a mais importante componente do Sistema Nervoso Central no processamento da memória, aprendizado e navegação espacial. Com o avanço da compreensão sobre seus circuitos e conexões internas, novos resultados e questionamentos encontram respaldo técnico, trazendo a construção de novas ferramentas mais acessíveis para visualização e compreensão dessas relações. O desenvolvimento de modelos computacionais simples, que ajudem na visualização da morfologia e dinâmica da atividade entre os principais núcleos e principais vias, é bastante importante no estudo dessa importante estrutura. Com isso, nesse trabalho, usando o software Neuronify(R), construímos um modelo hipocampal simplificado, mostrando as principais relações conhecidas entre as formações hipocampais (CA1, CA2, CA3 e CA4) e núcleos aferentes (Subiculum e Giro Dentado). Esse modelo mostra-se promissor principalmente como ferramenta de ensino.

Métodos: O modelo descreve os principais circuitos do hipocampo e suas correlações no processamento computacional da informação recebida. Utilizando-se da divisão funcional de subregiões da formação hipocampal em ‘*circuit motifs*’, foi estabelecido padrões de recursividade, inibição e propagação dos sinais de entrada em cada porção, delimitando as conexões entre tais microcircuitos. Utilizou-se uma representação com um número mínimo de neurônios, inibitórios e excitatórios, e sinapses para que a descrição fosse o mais intuitiva e ao mesmo tempo o mais morfologicamente real possível. Devido as limitações do simulador, neurônios glutamatérgicos e colinérgicos foram etiquetados como excitatórios e GABAérgicos como inibitórios, tal qual fibras musgosas, via perfurante e colaterais de Schaffer fossem adaptadas para uma visualização mais clara e objetiva.

Resultados: Foi construído um modelo computacional simples, usando a plataforma Neuronify(R). Nesse modelo usou-se 202 neurônios, inibitórios e excitatórios, conectados de acordo com a principal diagramação morfológica hipocampal. Esse arranjo permitiu boa compreensão da atividade que cada componente do circuito hipocampal apresenta, considerando suas conexões funcionais, e padrões de intensa inibição e disparos esparsos do giro dentado à excitação recorrente e coordenada observada no Cornu Ammonis (CA) 3. Desta forma, a atividade hipocampal pode ser vista através dos disparos dos neurônios e dos sensores disponibilizados pelo próprio simulador.

Discussão: O modelo proposto demonstrou ótima correspondência funcional com a dinâmica real hipocampal conhecida na literatura. Além disso, dispõe de diferentes possibilidades de análise dos sinais em cada nó da rede. Pode assim ser usado no ensino da dinâmica hipocampal ou mesmo servir como parâmetro de novas hipóteses acerca de distúrbios funcionais do hipocampo a partir de alterações em sua conectividade interna.

Agradecimentos: Esse trabalho foi parcialmente fomentado por CNPq, Plexon.

A mean-field model for conductance based integrate and fire neurons with variable timescales

Becker, Marcelo P.¹; Idiart, Marco A. P.¹;

¹Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, Brasil

Introduction: It is of great interest in computational neuroscience the ability to describe the firing rate of a neuron given a constant stimulus. For current-based integrate and fire neurons, this problem was solved by the work of Amit and Brunel¹ with a mean-field approach which became a standard in the field. Models of this type were improved to allow different variations of the integrate and fire neuron with more biological details. Here we propose a method for obtaining the firing rate of a conductance-based integrate and fire neuron with different timescales under constant stochastic input.

Methods: We construct a conductance-based integrate and fire model that receives Poisson-like excitatory and inhibitory inputs. Using a mean-field approximation, we reduce the system to a one-dimensional model with colored noise. Using the effective timescale approximation to eliminate the multiplicative part of the noise, we construct an effective Fokker-Plank equation with the Fox method. This allows us to use the results derived by Amit and Brunel¹ and calculate the firing rate of the model without free parameters. We compare the analytical rate results with simulations as a function of the excitatory time constant for different parameters of the neuron and the input.

Results: With the method utilized, we are able to obtain a good approximation for the firing rate in a large range of excitatory synaptic timescales. The model exhibits a sharp transition from low to high activity where the transition point depends lightly on the input firing rate (fig 1) but depends heavily on the synaptic weights. As expected, neurons under weaker input exhibit a slower transition to saturation, which also occurs for inhibition-dominated inputs. The error in the analytic results is larger in the region where the mean input is closer to the threshold, where fluctuations in the membrane potential become more important.

Discussion: Our model is capable of describing the firing rate of the biologically realistic integrate and fire neuron with a relatively simple approach and no free parameters. Most of the error appears in the region where fluctuations dominate, which suggests a problem with the effective timescale approximation. This assumption, however, can be dropped with the penalty of more involved calculations and less interpretable parameters. An extension in this way is currently under work.

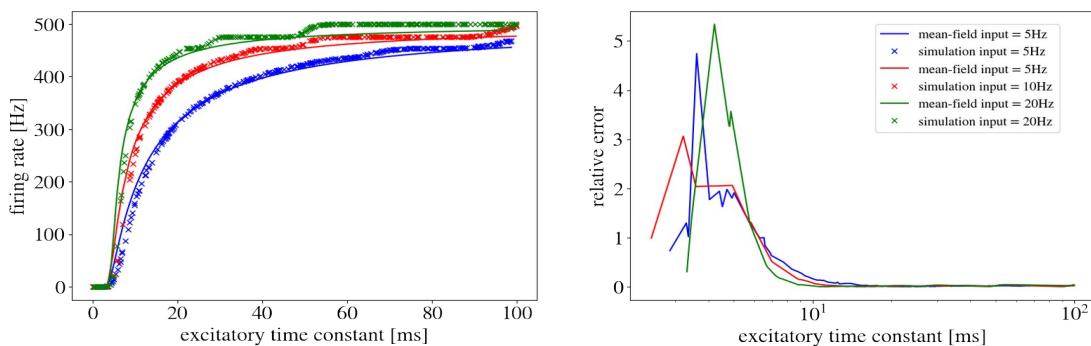


Fig. 1. (Left) Comparison of the firing rate obtained with the mean-field method (line) and simulations (cross) for 3 different input firing rate. (Right) Relative error between the model and simulations.

Acknowledgements: CNPq [131095/2020-0]

References: [1] D. J. Amit, N. Brunel, *Cereb. Cortex.* 7, 237–252 (1997).

Spatial Allocation of Memories in Modules using Theta-Gamma Traveling Waves

Gustavo D. Soroka¹; Marco Idiart¹

¹ Universidade Federal do Rio Grande do Sul, Rio Grande do Sul, Brasil

Introduction: Brain oscillations are fluctuations of the neuronal activity in the brain, often correlated with cognitive functions like reasoning, planning, language comprehension and memory. Despite being very well studied, the spatial properties of these rhythms are poorly elucidated due to experimental limitations, which recently have been overcome. Thereby, it is important to adapt working memory models based on oscillations since the evidences that these oscillations are in fact traveling waves are coming up. The research goal is to study the role of traveling waves in a multi-item working memory model using the phenomenon of memory allocation.

Methods: We propose a variation of the Lisman-Idiart model where the theta and gamma oscillations travel in space and the neurons are organized in modules. We constructed a network with four modules of integrate and fire current based neurons with the inhibitory-excitatory neuronal proportion of 1/5. Theta traveling waves in 8 Hz were modeled as a delayed oscillatory input current between the four modules and four different stimuli were given equally and sequentially to all modules in a gamma frequency. We vary the theta wave speed, gamma frequency of stimuli presentation and observed the process of memory allocation in the modules.

Results: We observed that exist an optimal gamma frequency of stimuli presentation (f_{γ}^{Best}) for each value of theta wave traveling speed (Ψ_{osc}) that allows exact one information to be allocated in each module in the correct sequence (Fig. 1). Replacing the theta wave for an alpha wave (12 Hz) impaired the allocation process.

Discussion: The implementation of a modular working memory model using traveling waves showed to be compatible with a modular network architecture. The theta traveling speed and gamma stimuli frequency related in a way to produce a wide range of values in which the network could perform the correct allocation of memories, showing itself as a flexible and stable phenomenon. The network performance impairment using frequencies in the alpha band is coherent with experimental evidences showing the inverse relation between alpha power and working memory performance. This work opens up the opportunity to explore spatial properties of brain oscillations in phenomena related to multi-items working memory.

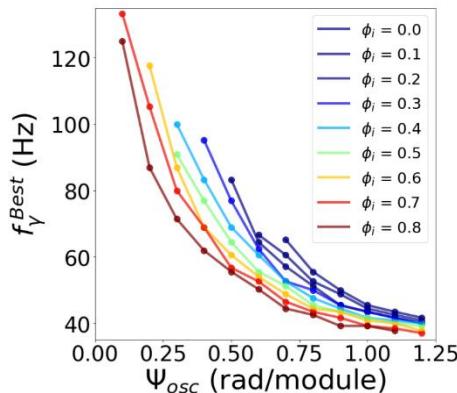


Fig. 1. Optimal gamma stimuli presentation frequency (f_{γ}^{Best}) for each theta wave travelling speed (Ψ_{osc}) in different initial time onsets (ϕ_i).

Acknowledgements: This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Izhikevich model simulation of spiking neurons in a Web-based architecture

Rodrigo de Sousa Gomide^{1,2}; Marcus Fraga Vieira¹

¹Federal University of Goiás, Goiânia, Brazil

²Federal Institute Goiano, Trindade, Brazil

Introduction: The development of mathematical models and the use of computer simulations has been widely used in the neuroscience field. These methods have made possible to conduct studies and discover new characteristics of the nervous system, which would hardly be possible using only experimental methods. The Izhikevich model was presented to reproduce the spiking and bursting behavior of known cortical neurons. This model presents an efficient computational performance, similar to the integrate-and-fire model. Furthermore, the Izhikevich model reproduce satisfactorily different firing patterns real biological neurons. This study presents a simulation environment of the Izhikevich model based on a WEB architecture. The objective of this study is to present a framework for an online simulation of a single neuron or a neuronal population according to the Izhikevich model.

Methods: The simulation environment was designed based on a client/server WEB structure. On the server-side, the model was developed in Python language using multiprocessing and advanced programming techniques (Cython). Also, a Rest API was implemented with three communication methods: GET (/); POST (/izhikevich and /izhikevichpool). In the application, on the client-side, the Angular Framework and Chart.js were used to generate the graphics, Fig. 1.

Results: The application allows the user to reproduce excitatory, inhibitory and other dynamic types of neurons proposed by Izhikevich (Fig. 1). It is also possible to test new types of neurons by individually changing the four internal parameters of Izhikevich model, and modifying the initial values of the membrane potential and of the membrane recovery variable. At the moment, the simulator only allows a constant step current as input.

Discussion: The simulator faithfully reproduces Izhikevich's formalism and allows the user to perform simulations of a single neuron or a population of excitatory and inhibitory neurons. The environment is responsive and allows the user to perform simulations on smartphones, helping the researcher in tests or studies in the neuroscience field.

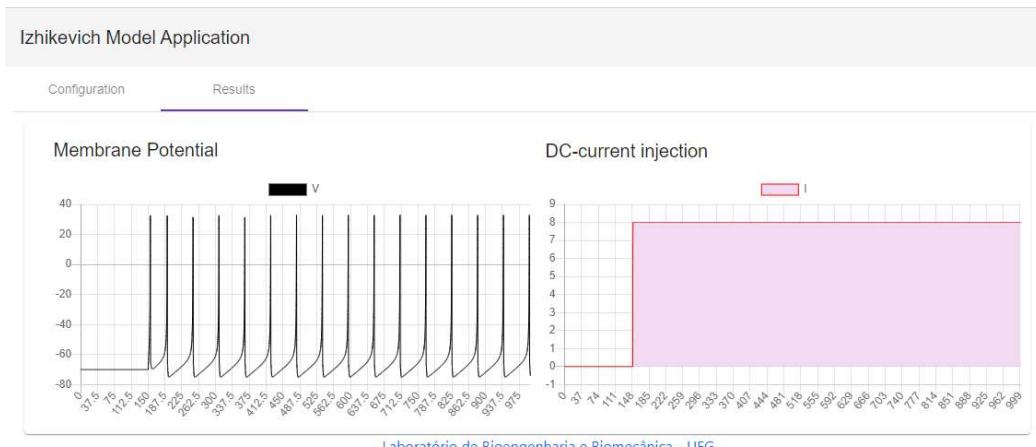


Fig. 1. Web simulation of the response of a simple excitatory cortical cell model (regular spiking)

Acknowledgements: There was no participation of a development agency.

Microscale modeling of seizures by spatiotemporal chaos in coupled map lattices

Rafael V. Stenzinger¹; Marcelo H. R. Tragtenberg¹

¹Universidade Federal de Santa Catarina, Florianópolis, Brasil

Introduction: Recent recordings at the microscale in animal models of epilepsy revealed some of the spatiotemporal characteristics of seizures. Pharmacologically reducing inhibition gives rise to a spontaneous dynamics of plane waves and transient spiral waves of electrical activity. Spirals in the biological tissue are traditionally associated with cardiac arrhythmias. Using the loss of synchronization of chaotic oscillators, previous reaction-diffusion models with conductance-based equations were proposed to simulate the formation of reentry and the initiation of the spiral rhythm in the heart. As described by these models and experiments, chaos may manifest in cardiac cells through abnormal oscillations in the action potential, known as Early Afterdepolarization (EAD). Inspired by these results, we generalize this approach to simulate the spiral formation in the brain. Our objectives are to evaluate whether the observed dynamics in epilepsy can arise from arrhythmia-inspired mechanisms and obtain insights about spiral formation in the brain.

Methods: We represent these phenomena in coupled map lattices using a three-dimensional map-based membrane potential model, the logistic KTz, capable of representing the behavior of healthy and unhealthy neurons and cardiac cells. Map-based models are far computationally faster than conductance models. We resort to computational simulations and tools from nonlinear dynamics, such as Lyapunov exponents and bifurcation diagrams, to generate and characterize our results. First nearest neighbors interactions are used in both cases. Cardiac simulations employ plateau spikes, diffusive couplings and stimuli delivered to one side of the network, simulating a pacemaker region. Neural simulations employ slow spikes, chemical synapses and a pacemaker neuron, simulating the epileptic focus.

Results: We use the KTz model to generalize previous results from the literature about cardiac reentry formation, such as finding chaos in the single cell and network, verify synchronization properties and spatiotemporal patterns. We show that chaos can manifest in the firing of neurons in our model through EAD-like oscillations, found experimentally in epileptic neurons. Spontaneous alternation of spirals with other oscillation rhythms is also observed in the network, similar to experiments. This alternation appears as changes in the amplitude and frequency of the network activity, qualitatively resembling electrographic seizures. Changes in direction and speed of the wave in relation to the pacemaker neuron are observed, similar to microelectrodes recordings close to the epileptic focus in humans.

Discussion: These results reinforce previous speculations about the relationship between arrhythmias and epilepsy and the question of chaos in the brain. We hypothesize that the disinhibited cortex acts similar to the excitable cardiac muscle and pacemaker neurons might dictate the oscillation rhythm, similar to ectopic pacemakers in the heart. This also suggests that heart and brain follow a reverse path in arrhythmia and epilepsy, with an increase of spatiotemporal complexity in the heart activity and a simplification in the brain.

Acknowledgements: R.V.S. thank financial support from FAPESC and CAPES. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Avaliação topológica sobre padrões espaço-temporais de mortalidade por COVID-19 em função de fatores socioeconômicos no estado de São Paulo

Rodrigo Lantyer¹, Priscila Antoneli¹, Marimélia Porcionatto¹, Jean Faber¹

¹Universidade Federal de São Paulo, São Paulo, BR.

Introdução: As medidas de prevenção e contenção não-farmacológicas da covid-19, como distanciamento, isolamento social e uso de máscara, mostraram-se efetivas, mas com limitações, principalmente em países com desigualdade social. Uma das grandes questões nessa pandemia é saber como essa desigualdade socioeconômica influencia as taxas de contaminação e mortalidade por covid-19. Porém, a heterogeneidade de fatores socioeconômicos, introduzem um grande problema na descrição da dinâmica de contaminação e mortalidade por covid-19; já que cada região apresentará tempos, necessidades e vantagens diferentes. Portanto, uma abordagem de enfrentamento homogênea sobre regiões heterogêneas, mas dependentes, provavelmente não será a mais adequada. Por isso, a construção de modelos baseados em dados reais, passível de atualização dinâmica, sobre a mortalidade da covid-19 e que considere a relação entre regiões e suas desigualdades sociais, pode ser crucial para políticas públicas mais efetivas.

Métodos: O modelo proposto considerou o estado de São Paulo como um modelo aproximado de avaliação nacional, devido a sua alta complexidade e heterogeneidade socioeconômica. Foi utilizado dados reais com 18 fatores socioeconômicos (tais como: vulnerabilidade social, moradias inadequadas, desemprego, renda, etc); dados sobre a taxa de mortalidade pela covid-19 nas 53 regiões geográficas imediatas de São Paulo, obtidos em diferentes bancos de dados. Aplicando-se a técnica multivariada PLS sobre as variáveis, $X = \text{'fatores socioeconômicos'}$ e $Y(t) = \text{'mortalidade quinzenal'}$, foi possível encontrar uma relação linear $Y = AX$. Tomando-se as distâncias euclidianas entre cada coordenada nas 3 primeiras componentes e um critério limiar para considerar apenas os padrões de similaridade mais próximos (<10%), foi possível a construção das matrizes adjacentes e os padrões de redes complexas, para cada quinzena ao longo de 503 dias de pandemia. Com os padrões de redes temporais, calculou-se 4 coeficientes topológicos associados às configurações globais das redes e 4 coeficientes topológicos associados a cada um dos 53 nós, em cada quinzena. Enfim, criou-se dois cenários futuros hipotéticos acrescentando diferentes padrões de mortalidade ao longo de 6 meses: (1) baixa mortalidade e (2) 3a. onda de mortalidade.

Resultados: Baseando-se na Teoria de Redes Complexas, construímos um modelo capaz de caracterizar a similaridade do histórico de mortalidade por covid-19, considerando os diferentes fatores socioeconômicos, entre as 53 regiões imediatas do Estado de SP. A partir desse modelo foi possível comparar os efeitos de cada índice topológico nos diferentes cenários. A relação entre os coeficientes topológicos e os fatores socioeconômicos indicaram que intervenções locais sobre regiões específicas podem contribuir de forma similar a intervenções globais.

Discussão: A técnica descrita buscou padronizar um método de análise quantitativa e qualitativa a partir das redes formadas, diferenciando o histórico de mortalidade por covid-19 a partir da topologia geral das redes e de fatores específicos de cada nó. O modelo mostrou-se robusto na descrição dos padrões topológicos ao longo do tempo e sensível na distinção dos cenários futuros, capaz de identificar regiões que desempenham papel crítico no controle na configuração das redes. Esse resultado sugere que estratégias de *lockdown* ou vacinação massiva dessas regiões se priorizadas poderiam gerar um efeito global similar a um enfrentamento homogêneo em todas as regiões.

Agradecimentos: CNPq, CAPES e FAPESP

Machine learning-based application for long-term electrocardiogram analysis

Juliana Mycaelle Silva^{1,2[0000-0002-6518-6892]}, Jonathan Araujo Queiroz^{1[0000-0001-8006-6242]}, Luís Fillype da Silva^{1[0000-0001-6689-8632]}, Gean C. Sousa^{1[0000-0003-0797-1099]}, Marta de Oliveira Barreiros^{1[0000-0002-1367-1968]}, and Allan Kardec Barros^{1[0000-0002-1654-0955]}

¹ Universidade Federal do Maranhão, São Luís, Brasil

² juliana_mycaelle@outlook.com

Abstract. Electrocardiogram (ECG) analyzes only can be performed by health professionals whose demand for care is often greater than the availability thus, this paper proposes the development of an application that seeks to facilitate the analysis for signs of ECG, seeking to make professional care more efficient. In this context, this work consists of the development of a application capable of processing long-lasting ECG signals to assist health professionals in making decisions. The application has an interactive interface that allows view the entire ECG signal in a single image generated by all overlapping cardiac cycles. The proposed application still has e-mail communication between users with the objective of facilitate patient follow-up. The application was tested on three different ECG signals, one artificial and two real, the first signal was an artificial signal generated in software Matlab, the second ECG signal has normal sinus rhythm, available in the MIT-BIH Normal Sinus Rhythm database. Database is the third ECG sign diagnosed with arrhythmia and can be found in the MIT-BIH Arrhythmia Database. The results obtained by the proposed method can be used to support decision-making in clinical practice.

Keywords: ECG, MATLAB, Application, Processing, Artificial Intelligence

1 Introduction

The heart is an organ that has cells with its own rhythm, capable of generating action potentials of which are recorded by the ECG, that is, the ECG is an exam that registers the variation of the electrical potentials of the cardiac muscle and is composed of the P wave that corresponds to the depolarization of the atria, the QRS complex that corresponds to the depolarization of the ventricles, and the T wave that records the repolarization of the ventricles [1].

Changes in heart rate patterns provide a sensitive and anticipated indicator of an individual's health compromises. High heart rate and signal good adaptation, characterizing a healthy individual with efficient autonomic mechanisms. On

the other hand, low heart rate and often an indicator abnormal or insufficient adaptation, which may indicate the presence of physiological malfunction in the individual, needing further investigations in order to find a specific diagnosis. This additional investigation can be performed based on the morphology of the cardiac cycles, deformations in the waves of the ECG signal, may be an indication of cardiovascular diseases [2].

Data released by the World Health Organization (WHO) indicate that 17.3 million people die each year worldwide victims of cardiovascular diseases, and 80% of these deaths are registered in low and middle income countries [3]. The number of deaths in the world is forecast to rise from 16.7 million in 2002 to 23.3 million deaths in 2030, which will put heart disease in a group responsible for 70% of all deaths in the world in 2030 [3]. In this context, this work proposes the development of an application that can assist in making decisions medical decision, processing long-lasting ECG signals. Thus, developing means to facilitate the diagnosis of cardiovascular diseases is essential in today's society, in this sense Gupta et al. (2010) proposes the development of an embedded system and GUI based on Matlab for ECG analysis acquisition [4]. In Ramkumar et al. (2021) developed a graphical user interface (GUI) design to display the detection of PQRST waves using the MATLAB tool where these peak values are plotted on the ECG signal waveform at each instant of its respective time [5].

For the development of an application that analyzes signals long-term ECG, we propose a generalization of the method presented in our previous study, Queiroz et. al [6], that is, in addition to the segmentation of the cardiac cycles proposed in our previous work, and of real importance in the analysis of the ECG signal, the segmentation of specific parts of the cardiac cycles such as, for example, the segmentation of the P wave, QRS complex and T wave. The segmentation of specific parts of the ECG signal are fundamental to assist the evaluation of the ECG by the specialist, for example the segmentation of the P wave can assist in the assessment of pathologies related to the atria, such as atrial fibrillation and atrial flutter [7, 8], since the segmentation of QRS complexes helps to evaluate pathologies related to ventricular depolarization and the T-wave segmentation helps in the evaluation of normalities associated with repolarization. [9]. Therefore, here we will present a method to segment the P waves, QRS complex and T wave of the ECG signal and not just the cardiac cycle as previously proposed [10, 11]. Subsequently, each segment of the ECG signal (cardiac cycle, P wave, QRS complex and T wave) is used to train three algorithms machine learning, which are: linear discriminant analysis, support vector machine and k-Nearest Neighbor. The results obtained are evaluated with the accuracy, sensitivity according to the performance of both algorithms in the test stage.

To achieve the usability of the proposed methodology, the graphic interface was created, for the direct use of the health professional. The purpose of the graphical interface is assist in monitoring long-term ECG signals reducing the analysis of the long-term ECG signal in a single image to be analyzed. The image presented by Graphical interface consists of the overlapping of all available car-

diac cycles of the ECG signal. The graphical interface still allows the sending of the processed signal and information added by the specialist, such as diagnostics via email. In this interface, we seek to assist the analysis of the long-term ECG.

2 Background

Usually, cardiac variability is used to provide temporal information of the ECG signal. The cardiac variability is obtained by the R-R interval, which is defined like

$$\text{interval} = R - R = RR = R_m - R_{m-1}, \quad (1)$$

m being the time point of the m -th peak R.

Several authors use the R-R interval, which is the difference between two consecutive R waves, as a source of data extraction. However, the analysis of RR intervals does not measure changes in other ECG signal waves, such as the P wave distortions for AF and the appearance of the F wave in the atrial flutter [7]. The proposed method in our previous study uses the voltage variability in each heartbeat, unlike the R-R interval, in which each cardiac cycle is associated with a single real number, the proposed method associates each cardiac cycle to a set of points, that is, to a vector. This method uses the variation of tension in each cardiac cycle is defined as

$$\mathbf{b} = (b_1, b_2, \dots, b_m), \quad (2)$$

where \mathbf{b} is a cardiac cycle, b_m the m -th sample in millivolt (mV) of \mathbf{b} with $L_I(\mathbf{b}) \leq m \leq L_S(\mathbf{b})$, in such a way that $L_S(\mathbf{b})$ is the upper limit and $L_I(\mathbf{b})$ the lower limit of m . The upper limit $L_S(\mathbf{b})$ and lower $L_I(\mathbf{b})$ are given by

$$L_S(\mathbf{b}) = P_R + F_s \lambda, \quad (3)$$

and

$$L_I(\mathbf{b}) = P_R - F_s \theta, \quad (4)$$

L_S and L_I being position limits, P_R the peak position R P_R is found using the Pan and Tompkins algorithm that has a 99.3% success rate of QRS complexes for the MIT-BIH arrhythmia database [1], F_s the frequency of sampling in seconds, λ and θ are heuristically assigned parameters and function as sliding windows in the cardiac cycle. Based on the duration of the P wave to Pico R(0.4s) and in the duration of peak R to T wave (0.6s), the values maximum values for healthy cardiac cycles should be $\theta = 0.4$ and $\lambda = 0.6$, that is, $\lambda + \theta \leq 1$ [5].

3 Materials and Methods

In this work, ECG data obtained through the following databases were used: MIT-BIH Normal Sinus Rhythm (NSR) and MIT-BIH Arrhythmia Database.

Each ECG signal was segmented into cardiac cycles of one second, characterized by 400 ms before and 600 ms after the R wave [4]. In addition, we propose here a generalization of our previous work, for that, we segment not only the cardiac cycles, but we will also present a technique to segment specific parts in the cycles cords, namely, segmentation of the P wave, segmentation of the QRS complex and segmentation of the T wave. Subsequently, each segment of the ECG signal (cardiac cycle, P wave , QRS complex and T wave) are used to train three machine learning algorithms: linear discriminant analysis, support vector machine and k-Nearest Neighbor. The results obtained are evaluated with the accuracy, sensitivity according to the performance of both algorithms in the test stage.

3.1 Database

The ECG signals used in this study were obtained from the public databases. Each of the databases corresponds to a class in the problem of discrimination of cardiac pathologies with the proposed methodology. The description for each database used is given as follows:

- MIT-BIH Normal Sinus Rhythm (NSR) database contains 18 ECG recordings of approximately 24 h duration. Subjects included in this database had no significant arrhythmias; they include 5 men, aged 26 to 45, and 13 women aged 20 to 50 [12];
- MIT-BIH Arrhythmia Database contains 48 half hour excerpts of two-channel ambulatory ECG recordings, obtained from 47 subjects [12];

ECG signal preprocessing At this stage, we used the algorithm of Pan and Tompkins. This algorithm was chosen because it does not require large computational effort and has a 99.3% success rate of the QRS complexes for the MIT-BIH Arrhythmia database [13]. The algorithm consists of a band-pass filter (Low Pass and High Pass Filters), derivatives, a squaring function, a moving window integration, threshold, and decision. The goal of preprocessing step is to reduce contamination of different types of noise and artifacts in the ECG signal.

Therefore, to perform this work, the following types of noise have been removed: a signal in the frequency of 60 Hz and its bandwidth below 1 Hz; baseline wander, a low-frequency (0.15 up to 0.3 Hz) noise that results from the patient inhaling and compels a baseline shifting of the ECG signals; electrode contact noise, that results from a deficiency in the contiguity between the electrode and skin, which adequately cuts off the measurement system from the subject; electrode motion artifacts, artifacts that result from variations in the electrode-skin impedance with electrode motion; muscle contractions, noise that results from the contraction of other muscles apart from the heart; electrosurgical noise, produced from other medical apparatus in the patient care circumstance at frequencies between 100 and 1 MHz; and instrumentation noise, noise produced by the electronic equipment utilized in the ECG measurements.

In the removal of the noise present in the ECG signal, the procedure specified by the methodology of [13] was used, where a bandpass filter is applied for the elimination of low and high noise. high frequencies, where in this task it was decided to use a Butterworth filter. When converting the ECG signal with the filter, high and low frequency noise will be eliminated. The applied filter has the shape shown in the 5.

$$\|H_n(j\omega)\| = \frac{1}{\sqrt{1 + (\omega)^{2n}}} \quad (5)$$

where H is the transfer function of the filter, j the number imaginary, n is the order of the selected filter, ω is the angular frequency of the signal in radians per second.

The amplitude of the ECG signal was normalized and the sampling frequency adjusted at 256 Hz with 12-bit resolution in a range of 10 milivolts. The start and end of each ECG signal record were excluded due to the measurement error characterized by the value of 1% of the signal duration.

3.2 Methods

The previous section presents an analysis for the cycle cardiac by the usual analysis and the analysis proposed by [6]. However, a cardiac cycle is composed of a set of ripples initiated by the sinoatrial module. Therefore, we propose use a generalization of the method proposed by Queiroz et. al for extracting data from each wave in the cardiac cycle. For this, it is necessary to rewrite the cardiac cycle (Eq. 6) as

$$C = P <> QRS <> T = \quad (6)$$

$$= (p_1, p_2, \dots, p_i) <> (qrs_1, qrs_2, \dots, qrs_j) <> \quad (7)$$

$$(t_1, t_2, \dots, t_k), \quad (8)$$

i, j and k being the lengths and $<>$ and the concatenation of the P , QRS and T waves, respectively, and $P \cap QRS \cap T = \{\}$.

When you want to isolate all waves from the ECG signal, upper and lower limits are defined for each wave. The lower limit of the P wave ($L_I(P)$) is given by Eq. 4 and the upper limit ($L_S(P)$) is given by

$$L_S(P) = \mathbf{t}_P F_s, \quad (9)$$

being the duration time in seconds of the \mathbf{t}_P wave.

The lower ($L_I(QRS)$) and upper ($L_S(QRS)$) limit of the QRS complex are given by

$$L_I(QRS) = L_s(P) + 1, \quad (10)$$

and

$$L_S(QRS) = \mathbf{t}_{QRS} F_s, \quad (11)$$

\mathbf{t}_{QRS} being the duration time in seconds of the complex QRS.

The lower limit ($L_I(T)$) and upper limit ($L_S(T)$) of the T wave are given by

$$L_I(T) = L_s(QRS) + 1, \quad (12)$$

and

$$L_S(T) = \mathbf{t}_T F_s, \quad (13)$$

\mathbf{t}_T being the duration time in seconds of the T wave.

3.3 Classifiers

The classification step consists of verifying the efficiency of the proposed methodology in segmenting the cardiac cycles, the P wave, the QRS complex and the T wave. This segmentation methodology is used to prepare the signal for the subsequent classification step in relation to the worked classes: some type of Atrial Fibrillation (FA) or Normal. For this, three machine learning algorithms are used to train and test: linear discriminant analysis, support vector machine and k-nearest neighbor.

- Linear Discriminant Analysis (LDA)

Discriminant analysis is a classification method. It assumes that different classes generate data based on different Gaussian distributions. Each class (Y) generates data (X) using a multivariate normal distribution. In other words, the model assumes X has a Gaussian mixture distribution.

For linear discriminant analysis, the model has the same covariance matrix for each class; only the means vary. To train (create) a classifier, the fitting function estimates the parameters of a Gaussian distribution for each class. To predict the classes of new data, the trained classifier finds the class with the smallest misclassification cost.

- Support vector machine

The support vector machine (SVM) aims to produce a binary classification model. Training vectors are mapped in a space of some dimension. Then SVM defines a separation hyperplane and its appropriate margin. For linearly separable data cases your linear kernel is adopted[14]:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j. \quad (14)$$

In cases where linear separation of samples is not possible, the use of the radial base function (RBF) kernel is required:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2), \gamma > 0. \quad (15)$$

This way the samples will be mapped nonlinearly in a larger dimension space.

- k-Nearest Neighbor

In the k-nearest neighbor (k-NN), in general, a memorization of the training set occurs. When a new sample is inserted, it will be labeled with its k nearest neighbors. The proximity of a new sample to its neighbors is defined by the Euclidean distance of its feature vectors [15]:

$$d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{l=1}^d (x_i^l - x_j^l)^2}. \quad (16)$$

- Multi Layer Perceptron

The neural networks Multi Layer Perceptron (MLP) have as their main feature the presence of at least one hidden neuron layer between the input layer and the output layer [16]. Thus an MLP network has at least two layers of neurons, distributed between intermediate layers and the output layer.

3.4 Evaluation Metrics

The results are evaluated in four categories, as follows: True positive (TP), in which a record with AF is classified with AF; True negative (TN), which in a normal record is classified as normal; False positive (FP) a normal record is classified with AF; False negative (FN), for this category, a record with AF is classified as normal. The methodology presented in the previous section is responsible for preparing the signal, using segmentation, for classification as FA or normal classes.

We use sensitivity (SENS), and accuracy (ACC) metrics to evaluate the classification stage. Sensitivity reflects how the adopted methodology is effective in correctly identifying the TP and Accuracy indicates how the method is effective in correctly performing the diagnosis.

The sensitivity, and accuracy are defined, respectively, by (17), and (18):

$$\text{SENS} = \frac{VP}{VP + FN} \times 100 \quad (17)$$

$$\text{ACC} = \frac{VP + VN}{VP + VN + FN + FP} \times 100 \quad (18)$$

To avoid overfitting, the cross-validation technique that evaluates the generalization capacity of a model will be used. Cross-validation partitions the dataset into mutually exclusive subsets, and later uses some of these subsets to estimation model parameters (training data), the remaining subsets (validation or test data) being employed in model validation. As a cross-validation technique, the K-fold method was used, which $K = 10$ [17].

3.5 Graphic interface

Where H is the transfer function of the filter, alrIn this work, the graphic interface that will be presented seeks to offer the user a facilitating means for the analysis of long-term ECG's, being an auxiliary means for the cardiology professional in the analysis and monitoring of patients. The interface is characterized by the easy use of the health professional, and the easy understanding, facilitating the analysis.

The proposed methodology was implemented in the numerical computation *Matlab* [18], and because they are algorithms implemented directly through a language programming, are not very easy to use for the end user. In order for the functions to have greater iterativity, it was decided to implement a graphical interface for the methodology proposed in the previous section.

The graphical interface is an application developed using the *APP Designer* available in the *Matlab* software version 2020b, *App Designer* is an interactive application development environment, where it is possible to design a layout for the application interface and program your behavior.

The platform is intuitive, offering the user an area to fill in information about the patient in question, and it is an environment where the end user, possible health professionals, such as cardiologists and professionals who need to perform ECG analysis, may have facilitated and streamlined the analysis of long-term ECG signals life in a simple way.

In the interface, the user uploads a file ECG signal in ".mat" format, and the application applies in this signal the proposed processing. Returning to the user at the interface the result of the ECG processing and the amount of Beats Per Minutes (BPM) of the signal analyzed, allowing the user an agility in the analysis long-term ECG's, making service of these professionals, as well as e-mail communication between health professionals or between professional and patient. Through of the interface, the user can save the graphics for analysis and send information you want to share via email.

Through App Designer, the implemented application can be shared as an installer file. In that case, it is necessary that the desktop on which the application will be used has installed the *Matlab* software, or sharing it by creating a Web application, where in this case the end user needs to have *MATLAB Compiler*, a compiler Matlab application that allows both sharing regarding the installation of web applications, installed on your desktop.

App Designer makes it possible to implement the application DAASINC. The interface was implemented to load and process an ECG signal previously saved by the professional of health, that when loaded in the Application is processed by the proposed method and the result of this processing.

4 Results and Discussion

The 14.575 cardiac cycles were obtained by applying methodology proposed in ECG signals obtained in banks Normal Sinus Rhythm MIT-BIH data [19, 12],

from where a signal with 4678 cardiac cycles was used, and arrhythmia MIT-BIH [20, 12], from where a signal with 6297 was used cardiac cycles, and by an artificial signal produced by a function implemented in Matlab with 3600 cardiac cycles. The classifiers used in this study were: Linear Discriminant Analysis (LDA)) with linear version; nearest k-neighbors (from english, k-Nearest Neighbors (k-NN)) with a range of 2-5 closest neighbors; support vectors machine (do English, Support Vector Machine (SVM)) with the polynomial nucleus. The 10-fold cross-validation technique was used to test classifiers. Overall performance of classifiers was assessed by taking the average accuracy of cross-validations. In addition, the parameters for the windows of the ECG signal are shown in Table 1:

Table 1: Parameters used to window the ECG signal ^{0.7}

| 3* Localization in the heart | 3* Terminology ECG | 2*Time [ms] | Limit | |
|---------------------------------|-----------------------|---------------|-----------|----------|
| | | | 2*Higher | 2*Bottom |
| Healthy Arrhythmia | | | | |
| Cycle cardiac | B | 1000 1500 | 500 80 | 0 |
| Atrium | P | 70-110 120 | 60 20 | 0 |
| Node atrioventricular | PQ | 120-200 | 110 25 | 20 |
| Depolarization Ventricular | QRS | 60-100 | 50 60 | 25 |
| Repolarization Ventricular | T | 90-130 | 70 80 | 65 |

Table II shows the sensitivity of the proposed method for ECG windows. The results are presented with a standard deviation of ± 3 samples.

0.8

Table 2: Sensitivity of the proposed methodology for winding the ECG signal.

| | *Sensitivity% | | | | | | | |
|------------|---------------|-----|-------|---------|---------|-------|-----|-------|
| | LS(P) | PP | Li(P) | LS(QRS) | Li(QRS) | LS(T) | PT | Li(T) |
| Healthy | 99.98 | 100 | 99.98 | 99.99 | 99.99 | 99.97 | 100 | 99.98 |
| Arrhythmia | 99.98 | 100 | 99.95 | 99.90 | 99.85 | 98.91 | 100 | 97.88 |

* Standard deviation of ± 3 samples; $L_S(P)$, $L_I(P)$ and P_P the upper and lower position limits and the peak wave P , respectively; $L_S(QRS)$ and $L_I(QRS)$ are the upper and lower position limits of the QRS complex, respectively; $L_S(T)$, $L_I(T)$ and P_T are the upper and lower position limits and the peak of the T wave, respectively.

The proposed method proved to be efficient in recognizing the beginning and the end of the windows in the ECG signal. For healthy patients, the average accuracy was greater than or equal to 99.97%. As for patients with arrhythmia, the best result found was 99.98% for detecting the beginning of the P wave and the worst result was for detecting the end of the T wave, with a hit rate of 97.88%. The high success rates of the model are due to the fact that the algorithm does not actually mark the signal, but it produces a window, with a search interval in the signal, from this search interval it is checked if it exists or not a window on the sign.

Table 3 shows the average accuracy of the classifiers (LDA, k-NN, SVM) for 5 different windows in the ECG signal (Cardiac cycles, P wave, QRS complex, T wave, segment PQ).

0.6

Table 3: ECG signal window using the proposed method in 5 different cases (Cardiac cycles, P wave, QRS complex, T wave, PQ segment).

| Average accuracy% | | | | | | | | |
|-------------------|-------|-------------|-------------------|-------|-------|-------|-------|-------|
| Training | Test | Classifiers | Cycles cardiac | P | QRS | T | PQ | |
| 3*Healthy | 3*18h | 3*4h | LDA | 99.99 | 99.98 | 99.98 | 99.98 | 99.90 |
| | | | k-NN | 99.99 | 99.99 | 99.99 | 99.98 | 99.92 |
| | | | SVM | 99.99 | 99.98 | 99.98 | 99.98 | 99.92 |
| 3*Arritmia | 3*18h | 3*4h | LDA | 99.97 | 99.52 | 99.57 | 99.03 | 99.02 |
| | | | k-NN | 99.97 | 99.97 | 99.97 | 99.95 | 99.80 |
| | | | SVM | 99.97 | 99.80 | 99.87 | 99.57 | 99.43 |

The classifiers were trained to recognize the different windows in the ECG signal (cardiac cycle, P wave, QRS complex, T wave, PQ segment). For pa-

tients healthy, the average accuracy was greater than or equal to 99.90%. This behavior is a consequence of the uniformity of the signal, in fact, there is little or no difference in signal morphology. The ECG signal with arrhythmia had average accuracy of up to 99.97% in heart rate classification cardiac wave P and QRS complexes; on the other hand, the lowest average accuracy was in segment segmentation PQ using LDA (99.02%). This is due to the fact that in some cardiopathies the contraction of the atria and ventricles occurs in a non-synchronized way, impairing the filling and ejection phases of the cardiac chambers, making it difficult thus, the distinction between atrial and ventricular electrical activity in the ECG signal [2]. ECG window was efficient in healthy heartbeat (Figura 2, campo Beat(B)) in all evaluated cases (heartbeat, P wave, QRS complex, T wave, PQ segment), and this behavior is a consequence of signal uniformity; indeed, there is little or no difference in signalmorpholog, differently from the heartbeat with arrhythmia (Figure 5 field Beats(B)). P wave with arrhythmia (Figure 5 field Beats(P)) present deformations compared to healthy P wave, a consequence of the arrhythmia's associated with atrium (e.g., atrial fibrillation, atrial flutter) [9, 21]. Ventricular diseases, such as heart block, cause irregularities in the electrical activity of the ventricles, which is evident in QRS complexes with arrhythmia (Figure 5), a characteristic not observed in healthy patients (Figure 4) [22, 23]. The morphology of the T wave is completely apparent in healthy heartbeats; however, in heartbeats with arrhythmia, the T wave is deformed [10, 11]. The PQ segment is completely visible in healthy heartbeats (Figure 4); nevertheless, due to increased heart rate, the PQ segments are reduced in heartbeats with arrhythmia (Figure 5), in some cases a consequence of the AV blocked [24].

The implementation of the method the proposed strategy proves to be relevant, above all, in low and medium income, where 80% of deaths from arrhythmias occur [2], or even in countries of great territorial extension, such as Brazil, in which the concentration of the number of doctors and 2.09 doctors per 1000 in habitants. In this context the application is presented to assist the health professionals in the analysis of long-term ECG signs duration, the (DAASINC). The aim of DAASINC is to assist the assessment of cardiac condition using the methodology proposed in [4]. The test application is available at the link: <https://1drv.ms/u/s!AsvmLwSVT8QIaucopyXefyTKjk?e=fNkbaX> and for its execution it is necessary that the *Matlab* software installed on the desktop that the user uses, also if makes it necessary that the signal to be processed is in the same application workbook.

In the implementation of the graphical interface so that the user can process the signal directly through the application, the interface shown in Figure 1.

In Figure 1, we can see, in general, the interface graphics of the application, where we note the interactivity allowing the user a skillful manipulation. Initially, right we observe the instructions area of the application, where we see the area for inserting the data on the ECG that will be analyzed. After requesting the ECG to be processed by the user select in the check box if you want to save the figures resulting from processing. Just below, after processing the signal, the amount of

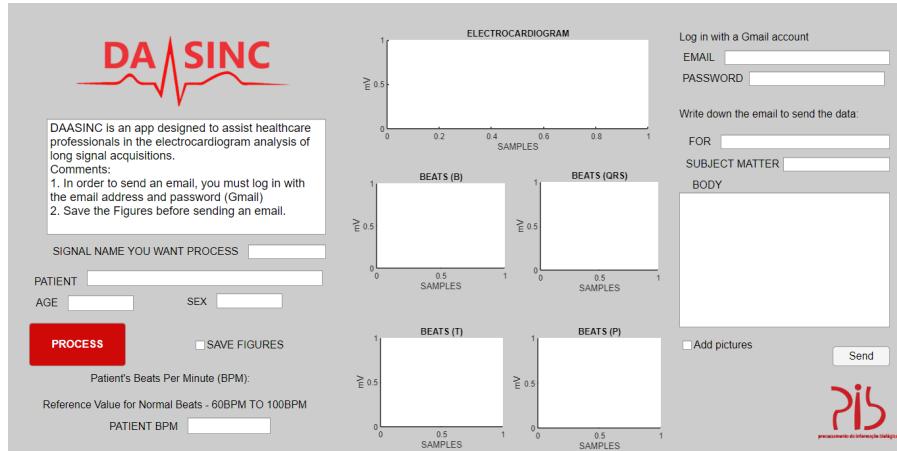


Fig. 1: DAASINC Graphical Interface - An interface contains three areas, the first on the right, contains instructions for use and an area for filling in patient data, it also contains an area for displaying an amount of BPM. In the second area, a central area, we see the displays for plotting the wave graphs of the processed ECG signal. In the third area, on the left, we see the area for sending e-mails.

BPM will be shown on the application screen, which can be compared with the limit values and can be used by the professional to help classify the signal. In the central area of the interface we observe the plotting areas graphs related to the ECG, where in the first figure and the signal shown as previously collected by the user, and in the following figures the processed signal of the cycles P waves, QRS complexes, and T waves. All the figures show a set of suspended controls that through them the user can increase or decrease the visualization of the ECG signal, as desired.

You can also see in the interface the area for sending e-mails, on the left of Figure 1, where the user after log in to a google server account, you can send an email to any user with analysis and comments. In this area, we note that by definition, the subject of the e-mail contains the patient's name, and the body of the email attached in its structure to the end of the typed text by the user the information of age, sex and quantity of User Beats Per Minute (BPM). Still, in the area of sending e-mails we observe the check box destined to the choice of the user about the attachment of the images of the processing of the ECG in email.

When the interface is used, Figure 2 is obtained, where it is possible to observe in the first area of the interface the filling of data around the processed ECG signal, this filling done by the end user, just below the press button to process the loaded ECG signal by the user, we observe the amount of BPM identified

Machine learning-based application for long-term electrocardiogram analysis

by the signal processing. In area 2 of the Figure we observe the graphs resulting from the processing of a normal ECG signal through the interface, we observe that in the interface the user will be able to approximate the visualization of the ECG signal or the desired segment, through the zoom function in the desired area of the graphs. In area 3 of the Figure we can observe the area for sending the e-mail filled with the sending data by the end user, when pressing the send button, the written e-mail will be sent and attached The figures saved in the execution of the interface.

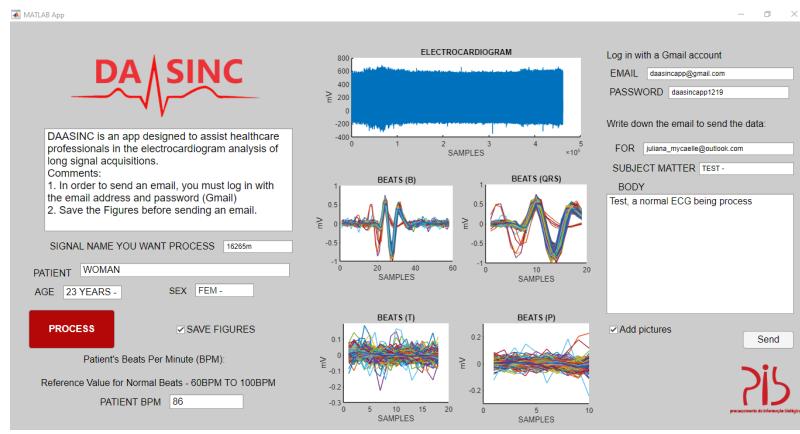


Fig. 2: DAASING interface Working - The interface in its normal operation, where we can see in the first two areas the result of processing a normal ECG signal. We can see the graphs plotted by the processing of the ECG signal and the information filled around the signal by the end user (Fictitious information).

0

When submitting the application to tests, it was decided to process three distinct signals with the same duration, 1 hour of duration each signal. The first one, and an artificial signal, and the processing of this signal in the application is shown in Figure 3. Then a normal ECG signal was processed, where the the

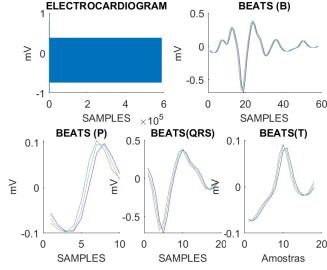


Fig. 3: Artificial ECG signal processed in DAASINC - ECG signal of 3600 cardiac cycles processed by the application, in the first graph the electrocardiogram in its full extension, the user can zoom in on the desired areas in the application. In the following graphs we see the processed graphs of B, P, QRS and T, we see few lines because it is an artificial signal, clean, without noise.

result of this processing is shown in Figure 4.

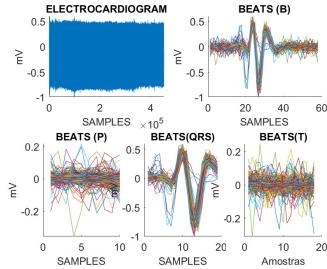


Fig. 4: Normal ECG signal processed in DAASINC - In the first graph the complete ECG, the user can zoom in on the desired areas in the application. In the following graphs we see the processed graphs of B, P, QRS and T, we can see many overlapping lines indicating a normal signal.

Soon after, in Figure 5, we observe the third sign processed, a signal with cardiac arrhythmia was chosen.

When we look at Figure 3, we notice the occurrence of few lines in the image, which was to be expected, since that an artificial signal always has the same beat,

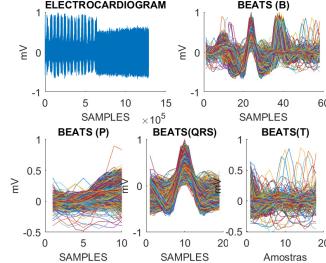


Fig. 5: ECG signal with arrhythmia processed in DAASINC - In the first graph the complete ECG, the user can zoom in on the desired areas in the application. In the following graphs we see the processed graphs of B, P, QRS and T, we can observe a disorder in the processed lines, indicating the arrhythmia.

without presenting the variations normally found in a real ECG signal. In Figure 4, we observe the variations intrinsic to the ECG, with variations in amplitudes and lags. In Figure 5, when we observe an ECG signal with arrhythmia, we notice the characteristic disorder of this signal. Comparing three signs, we observed a progression in the disorder starting of the artificial ECG signal as the most ordered, passing by the healthy signal until the signal with arrhythmia that presents greater disorder, indicating the progression in the alterations of the standard characteristics of the ECG signal.

Also, at the end of the first application area, the amount of BPM of the analyzed signal is observed after processing. In Figure 6, we note the amount of BPM for the signal Normal database used in the tests.

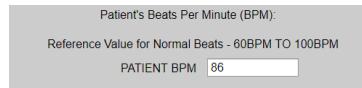


Fig. 6: Amount of BPM for a normal ECG processed in DAASINC - Amount of BPM for a patient with a normal ECG signal processed by the application, the displayed value can be compared with the limits shown in the application.

In Figure 6, we can see the limits for a BPM to be considered normal, and we observe the processing result for the normal ECG signal used for testing the app. Through this result shown in the interface, the user may have more information at his disposal for medical decision making.

In the section for sending email, the user can make login with a Gmail account and password, and just below insert the recipient and the subject of the email. In the field immediately below, the user can write information about the signal

analyzed, be it information for patients or information to be shared with other health professionals.

Then the user decides on the attachment of the figures of ECG processing and chooses to send the email. In the figure 7 we can notice the result of an email received through application.

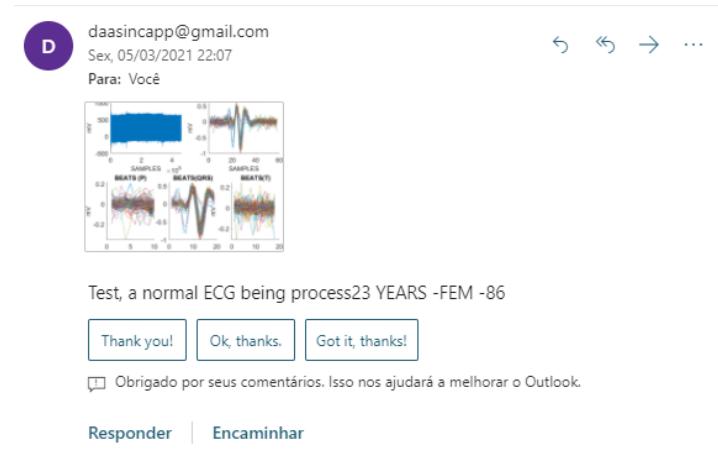


Fig. 7: E-mail sent using the DAASINC interface - An e-mail was sent with an ECG from a 23-year-old female patient whose beats were from a normal ECG signal, the data from the ECG signal shown are fictitious.

When sending the email, the user will be able to communicate remotely with other users of the application, being able to be patients or other health professionals, which would facilitate distance service, in places with difficult access type of iteration is important for monitoring cardiac patients.

5 Conclusion

In practice, develop a computational method to aid diagnosis for cardiac pathologies using the ECG has at least two major challenges to be overcome. One is the computational cost and the other is the infeasibility of obtaining the ECG signal for long periods of time to that the algorithms can be trained and tested. In order to overcome this challenge in parts, a method was presented able to analyze the ECG signal in a systematic and simplified way. The simplicity of the proposed method in segmenting cardiac cycles and subsequently each wave of the same cardiac cycle, made it possible to implement an interface graph that can be used to aid diagnosis applications in systems similar to the 24-hour Holter. The DAASINC application facilitates the care and monitoring of patients

whose physical distances make it difficult to access by health professionals, and the patient's ECG analysis can be streamlined through the interface and shared with more Health professionals. The proposed methodology has been tested on long-term ECGs, however, it can be used to assist in the evaluation of any ECG signal. So our next challenge and evaluate the proposed methodology in real time.

This work represents an important step in the development of an aid system for the medical diagnosis of cardiovascular diseases to be developed based on this initial step. In addition to representing an academic research aid system, it also represents the beginning of a facilitation of the analysis of ECG exams. The main contribution of this interface is the initial step towards the greater objective of helping to monitor cardiovascular diseases.

References

1. Hall, John E. 2015. Guyton and Hall Textbook of Medical Physiology. 13th ed. Guyton Physiology. London, England: W B Saunders.
2. Luiz Carlos Marques VANDERLEI1, Carlos Marcelo PASTRE, Rosângela Akemi HOSHI, Tatiana Dias de CARVALHO, Moacir Fernandes de GODOY, Basic notions of heart rate variability and its clinical applicability, Rev Bras Cir Cardiovasc 2009; 24(2): 205-217.
3. WHO; World Heart Federation; World Stroke Organization, 1st ed. 164 p. 2011. ISBN: 978 92 4 156437 3.
4. GUPTA, R.; BERNA, J. N.; MITRA, M. Development of an embedded system and MATLAB-based GUI for online acquisition and analysis of ECG signal. *Measurement*, v. 43, n. 9, p. 1119-1126, 2010.
5. RAMKUMAR, M. et al. A Graphical User Interface Based Heart Rate Monitoring Process and Detection of PQRST Peaks from ECG Signal. In: *Inventive Computation and Information Technologies*. Springer, Singapore, 2021. p. 481-496.
6. Queiroz, J., Azoubel, L. Barros, A. Support system for classification of beat-to-beat arrhythmia based on variability and morphology of Electrocardiogram. EURASIP J. Adv. Signal Process. 2019, 16 (2019). Available: <https://doi.org/10.1186/s13634-019-0613-9>
7. A. d. P. Mansur and D. A. Favarato, "Trends in Mortality Rate from Cardiovascular Disease in Brazil, 1980-2012," Arquivos Brasileiros de Cardiologia, vol. 107, pp. 20 – 25, 07 2016. [Online]. Available: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2016004000020nrm=iso
8. Augustin J. DeLago, Mohammed Essa, Alireza Ghajar, Matthew Hammond-Haley, Arshi Parvez, Isaiah Nawaz, Joseph Shalhoub, Dominic C. Marshall, Saman Nazarian, Hugh Calkins, Justin D. Salciccioli, Binu Philips, Incidence and Mortality trends of Atrial Fibrillation/Atrial Flutter in the United States 1990 to 2017, The American Journal of Cardiology, 2021, ISSN 0002-9149, Available <https://doi.org/10.1016/j.amjcard.2021.02.014>.
9. M. Masè, M. Disertori, M. Marini, F. Ravelli, Characterization of rate and regularity of ventricular response during atrial tachyarrhythmias. insight on atrial and nodal determinants. *Physiol. Meas.* 38(5), 800 (2017).

10. G. T. Lines, B. L. de Oliveira, O. Skavhaug, M. M. Maleckar, Simple t-wave metrics may better predict early ischemia as compared to st segment. *IEEE Trans. Biomed. Eng.* 64(6), 1305–1309 (2017). Available: <https://doi.org/10.1109/TBME.2016.2600198>
11. J. Ramírez, M. Orini, J. D. Tucker, E. Pueyo, P. Laguna, Variability of ventricular repolarization dispersion quantified by time-warping the morphology of the t-waves. *IEEE Trans. Biomed. Eng.* 64(7), 1619–1630 (2017). Available: <https://doi.org/10.1109/TBME.2016.2614899>
12. PHYSIONET (ed.). PhysioBankATM. [S. l.]: PhysioNet, 2021. Available: <https://archive.physionet.org/cgi-bin/atm/ATM>
13. PAN J.; TOMPKINS, W. J. A real-time qrs detection algorithm. *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING*, v. 32, n. 3, p. 230–236, 1985.
14. Evgeniou T., Pontil M. (2001) Support Vector Machines: Theory and Applications. In: Palioras G., Karkaletsis V., Spyropoulos C.D. (eds) Machine Learning and Its Applications. ACAI 1999. Lecture Notes in Computer Science, vol 2049. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-44673-7_12
15. Cunningham, Padraig Delany, Sarah. (2007). k-Nearest neighbour classifiers. *Mult Classif Syst.*
16. Fionn Murtagh, Multilayer perceptrons for classification and regression, *Neurocomputing*, Volume 2, Issues 5–6, 1991, Pages 183-197, ISSN 0925-2312, Available: [https://doi.org/10.1016/0925-2312\(91\)90023-5](https://doi.org/10.1016/0925-2312(91)90023-5).
17. Hossin, Mohammad M.N., Sulaiman. (2015). A Review on Evaluation Metrics for Data Classification Evaluations. *International Journal of Data Mining Knowledge Management Process.* 5. 01-11. Available: <https://www.aircconline.com/ijdkp/V5N2/5215ijdkp01.pdf>
18. MATHWORKS (United States). MATLAB. In: Copyspider: Software Anti-plágio. Version 2020b. [S.I.], Available: <https://ch.mathworks.com/d>.
19. R. A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. 215–220, 2000.
20. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, physiotoolkit, and physionet,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000. [Online]. Available: <http://circ.ahajournals.org/content/101/23/e2157>
21. Y. Huo, F. Holmqvist, J. Carlson, T. Gaspar, G. Hindricks, C. Piorkowski, A. Bollmann, P. G. Platonov, Variability of p-wave morphology predicts the outcome of circumferential pulmonary vein isolation in patients with recurrent atrial fibrillation. *J. Electrocardiol.* 48(2), 218–225 (2015). Available: <https://doi.org/10.1016/j.jelectrocard.2014.11.011>.
22. D. C. Sigg, P. A. Iaizzo, B. He, *Cardiac Electrophysiology Methods and Models*. (Springer, New York, 2010). ISBN 978-1-4419-6657-5 e-ISBN 978-1-4419-6658-2 Available: <https://doi.org/10.1007/978-1-4419-6658-2>
23. S. Lee, D. Park, K. H. Park, Qrs complex detection based on primitive. *J. Commun. Netw.* 19(5), 442–450 (2017). Available: <https://doi.org/10.1109/JCN.2017.000076>
24. B. M. Glover, P. Brugada, *Clinical Handbook of Cardiac Electrophysiology*. (Springer International Publishing, Brussels, 2016). Available: <https://doi.org/10.1007/978-3-319-40818-7>

Modelo baseado em aprendizado de máquinas para classificar atividades encefálicas cognitivas *

Juliana Mycaelle Silva^{1,2[0000-0002-6518-6892]}, Gean C. Sousa^{1[0000-0003-0797-1099]}, George V. S. Souza^{1[0000-0003-4496-834X]}, Jonathan A. Queiroz^{1[0000-0001-8006-6242]}, and Allan Kardec Barros^{1[0000-0002-1654-0955]}

¹ Universidade Federal do Maranhão, São Luís, Brasil

² juliana_mycaelle@outlook.com

Abstract. Os diferentes tipos de atividades realizadas pelo sujeito, como ler; ouvir música; dançar; entre outras, conduzem a ativação das regiões encefálicas. Dentre essas atividades, as atividades cognitivas estão associadas a uma ativação de regiões encefálicas relacionadas com a aprendizagem, como as regiões do lóbulo frontal (Superior Frontal, Precuneus, entre outras). Diversos estudos vêm sendo desenvolvidos para relacionar atividades cognitivas e as regiões encefálicas associadas. Este tipo de estudo é importante na compreensão da funcionalidade e conectividade do encéfalo e esse conhecimento pode servir de auxílio para o diagnóstico de anormalidades no seu funcionamento. Este estudo tem como objetivo a elaboração de um modelo classificador de atividades encefálicas durante o desenvolvimento de uma atividade cognitiva dentro das três categorias de atividade: Jogo de video game, música ou matemática. Para isso, utilizou-se dos sinais de Eletroencefalograma (EEG) coletados em duas bases de dados públicas e com uso da técnica de estimativa de fontes encefálicas estimou-se as regiões anatômicas relacionadas a cada uma das atividades. Após essa determinação, foi treinado e testado um modelo baseado em aprendizado de máquinas que classifica o tipo de atividade desempenhada de acordo com as categorias de atividade. Dos resultados obtidos, podemos ressaltar o modelo classificador das atividades encefálicas elaborado com acurácia de 99,9%. Este trabalho se faz um passo importante na elaboração de um sistema de monitoramento de sinais de EEG considerando a ativação encefálica.

Keywords: Eletroencefalograma, EEG, Aprendizado de máquinas, Inteligência Artificial.

1 Introdução

O encéfalo é o principal órgão do sistema nervoso central [1] e a funcionalidade de cada região do encéfalo pode ser estudada para fornecer informações a cerca do desempenho nas atividades desenvolvidas pelo indivíduo [2, 1]. O conhecimento

* Este trabalho foi financiado por FAPEMA (75121/2021).

de seu funcionamento pode servir de auxílio para o diagnóstico de anormalidades de cunho patológico, fisiológico e funcionais. Essas anormalidades podem gerar deficiências ou problemas que afetam a capacidade cognitiva do indivíduo [1].

Dentre as atividades desenvolvidas pelo indivíduo, as atividades cognitivas estão associadas a uma ativação de regiões encefálicas relacionadas com a aprendizagem, como as regiões do lóbulo frontal [3, 4]. Sabendo que essas atividades geram uma ativação de regiões encefálicas específicas nas distintas atividades cognitivas[5], o objetivo desse trabalho é desenvolver um modelo para classificação da atividade neuronal relacionada aos estímulos emitidos durante as atividades cognitivas de jogar video game, responder questões de matemática ou ouvir música. Esses estímulos emitidos durante essas atividades estão relacionados à concentração, emoção e habilidade [6–9].

Para avaliar nossa hipótese utilizaremos a base de dados produzida no trabalho de Cavanagh et al. (2021), cuja base de dados é disponibilizada na plataforma *Openeuro* [10], nele foi realizada uma investigação para compreender o significado psicológico dos eventos cerebrais ocorridos durante o jogo de video game onde o EEG foi tratado através da classificação de padrões de frequência na superfície do couro cabeludo, os chamados topomaps [4]. E a base de dados produzida no trabalho de Duan et al. (2021), cuja base de dados também é disponibilizada na plataforma *Openeuro* [11], onde foi realizada uma análise para observar as potências da atividade encefálica em cinco situações: abrindo os olhos, fechando os olhos, ouvindo música, respondendo questões matemáticas, teste de memória. Nele o EEG foi tratado através da análise de potência na superfície do couro cabeludo [12].

No que tange à funcionalidade do encéfalo durante a realização de atividades cognitivas, Zhang et al. (2021) realizou uma análise exploratória das mudanças relacionadas à idade na ativação do córtex encefálico e as condições funcionais do órgão durante um jogo de video game, no trabalho, eles buscaram uma associação entre as regiões ativadas e a atividade de jogar video game, realizando uma comparação entre idosos e jovens [6]. Liu et al. (2020) realizou uma análise do comportamento da ativação encefálica durante a audição de cinco ritmos musicais distintos, procurando encontrar uma relação entre ritmo musical e o tipo de ativação apresentado. Ainda foi destacado diferenças na ativação encefálica devido aos diferentes estilos musicais [9]. Wilkey et al. (2017) investigaram a ativação neural infantil associada a aprendizagem de questões aritméticas. Neste experimento eles observaram relação entre a atividade elétromagnética do lóbulo parietal, frontal e occipital com a atividade de responder questões matemáticas [13]. Em todos estes trabalhos citados, a tecnologia de captação de imagem da atividade cerebral utilizada foi a ressonância magnética funcional (fMRI). Essa tecnologia é de alto custo e sua manutenção e seu uso requerem um alto conhecimento [2]. Portanto, deixaremos nesse trabalho de utilizar o fMRI como alternativa padrão e usaremos o EEG, pois o exame de EEG tem custos menores que o exame fMRI e por esse motivo é de fácil acesso. A substituição da fMRI pelo EEG nem sempre é possível, mas, com a técnica utilizada neste trabalho,

os sinais de EEG foram suficientes para determinar as regiões encefálicas relacionadas as atividades desenvolvidas pelos indivíduos.

Neste trabalho, a classificação das atividades cognitivas “Jogo de video game, música, matemática”, será realizada aplicando-se a técnica de estimativa de fontes encefálicas nos sinais de EEG coletados em [4] e [12], por meio dessa técnica é possível identificar regiões encefálicas ativas, ou seja, que apresenta alguma atividade eletromagnética durante a atividade desenvolvida pelo indivíduo. Partindo das regiões encefálicas associadas a cada atividade, supomos que seja possível desenvolver um modelo que classifique a atividade desempenhada dentre as classes sugeridas, para isso categorizamos as atividades em três grupos “Jogo de video game, Música, Matemática”.

2 Metodologia

Para o desenvolvimento do modelo, inicialmente foi realizada uma busca pelas bases de dados dos sinais de EEG que seriam utilizados. Após a busca pelas bases de dados foi realizado a estimativa das fontes encefálicas, onde foram encontrada as regiões anatômicas emissoras dos sinais de EEG analisados. A partir das regiões determinadas, fora realizada a elaboração de um banco de dados contendo todas as regiões encefálicas associadas a cada uma das atividades “Jogo de video game, Música, Matemática”. A partir do banco de dados criado, obteve-se o vetor de características necessário para o treinamento e teste do modelo, onde as características do vetor de entrada foram as regiões encefálicas relacionadas com cada uma das atividades. O passo seguinte foi a elaboração do modelo classificador, nesta etapa, foram testados alguns modelos de classificadores baseados em aprendizado de máquinas, e a partir dos modelos testados foi eleito um modelo para uso, o modelo escolhido foi treinado e testado. As etapas podem ser vistas de maneira visual na Figura 1.

Seguindo-se as etapas apresentadas, será explanado a seguir detalhadamente o que foi realizado em cada etapa, detalhando cada uma delas.

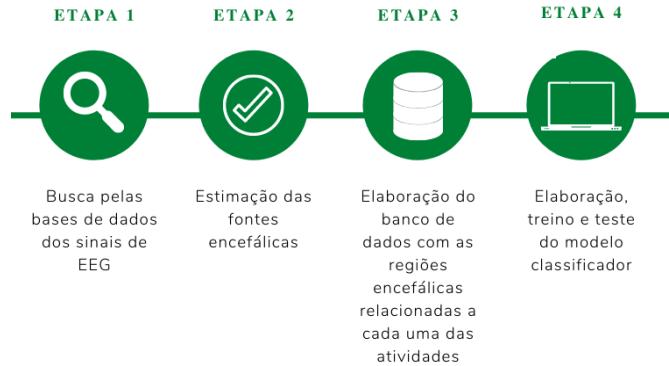


Fig. 1: Passos utilizados na elaboração do modelo classificador. Na etapa 1 foi realizada a busca pelas bases de dados dos sinais de EEG. Na etapa dois foi realizada a estimação de fontes encefálicas. Na etapa foi elaborado um banco de dados com as regiões encefálicas associadas a cada uma das atividades. Na etapa quatro foi elaborado, treinado e testado o modelo classificador.

2.1 Etapa 1 - As bases de Dados

As bases de dados utilizadas foram as bases de dados ds003517 utilizada em [4] e disponível na plataforma *Openneuro* [14]¹ e a base de dados ds003685 utilizada em [12] disponível na plataforma *Openneuro* [11]².

Na base ds003517, está registrado o EEG de 17 indivíduos normais com idade média de 20,94 anos. Dos 17 participantes 11 são do sexo masculino e não apresentam nenhum tipo de disfunção ou problema de desempenho em atividades cognitivas. Os sinais de EEG da base de dados foram registrados continuamente em 0,01 a 100 Hz com uma taxa de amostragem de 500 Hz e uma referência CPz online em um sistema Brain Vision de 64 canais durante o jogo contínuo de um video game onde cada tarefa envia gatilhos para o arquivo EEG e também produz dados contínuos em um arquivo de registro [14]. Os dados foram coletados por volta de 2015 no Laboratório de Ritmos Cognitivos e Computação da Universidade do Novo México [14].

Na base ds003685, está registrado o EEG de 57 participantes saudáveis dos quais 24 são homens, com idade média de 19,7 anos, que não apresentam nenhum tipo de disfunção ou problema de desempenho em atividades cognitivas. Os indivíduos realizam algumas tarefas: Abrir os olhos, fechar os olhos, responder questões de matemática, ouvir música, teste de memória [12]. Os sinais de EEG da base de dados foram registrados continuamente com uma taxa de amostragem de 500 Hz e uma referência CPz online em um sistema EEGLAB de 63 canais,

¹ <https://openneuro.org/datasets/ds003517/versions/1.1.0>

² <https://openneuro.org/datasets/ds003685/versions/1.0.2>

para este trabalho foi selecionado os sinais de EEG coletados durante as seções em que os participantes responderam questões matemáticas e ouviram músicas.

2.2 Etapa 2 - A estimação das fontes encefálicas

Neste trabalho propomos o uso dos sinais de Eletroencefalograma (EEG) como forma de obtenção de imagens da atividade encefálica, pois por meio da aplicação de uma técnica conhecida como estimativa de fontes encefálicas no EEG, é possível obter as regiões anatômicas do encéfalo relacionadas com o sinal de EEG capitado [2]. A estimativa de fontes encefálicas consiste em um mecanismo que, dado o sinal de EEG ou uma imagem da atividade elétromagnética encefálica, é capaz de identificar as regiões do encéfalo que são ativadas durante um evento qualquer [2, 1].

Nesta etapa utilizou-se o software de análise de sinais neurais Brainstorm [10]. O software tem entre outras funções, a função de realizar a estimativa de fontes encefálicas, neste trabalho em específico utilizou-se o algoritmo sLORETA para a estimativa das fontes encefálicas [15], a partir do algoritmo é realizada a estimativa das fontes no software. A estimativa das fontes encefálicas foram necessárias para identificar a região encefálica relacionada com o sinal de EEG de cada uma das atividades “Jogo de video game, Música, Matemática”.

2.3 Etapa 3 - O banco de dados

Após a estimativa das fontes encefálicas associadas a cada uma das atividades, fez-se uma contagem das regiões anatômicas associadas ao EEG emitido durante a atividade, nesta etapa considerou-se o atlas anatômico *Mindboggling* [16]. O atlas *Mindboggling* possui 29 regiões anatômicas definidas, e foi desenvolvido visando o auxílio da identificação de doenças mentais através da localização de regiões de interesse em exames de imagem encefálicas[16].

A contagem das regiões foi feita através da anotação da ativação de cada região relacionada a cada uma das atividades, onde adotou-se a seguinte metodologia: Caso a área γ tenha sido ativada na atividade Θ , esta área receberia 1, caso contrário receberia 0, para este trabalho é considerado região ativa a região do encéfalo que apresenta alguma atividade elétromagnética encefálica durante o evento específico, essa atividade apresenta-se como uma corrente elétromagnética na região encefálica diferente de zero e é observada na etapa de estimativa de fontes encefálicas.

A partir da elaboração do banco de dados foi feito o vetor de características que seria entrada para o modelo classificador, o vetor possuía a seguinte forma:

$$X = \begin{bmatrix} x_{1,1} & \dots & x_{1,29} \\ \vdots & & \vdots \\ x_{n,1} & \dots & x_{n,29} \end{bmatrix} \quad (1)$$

onde cada x é uma região encefálica que está associada a cada uma das atividades “Jogo de video game, Música, Matemática”. A dimensão do vetor de características é 29 porque existem 29 regiões anatômicas no encéfalo anatomicamente.

Como as bases de dados possuem quantidade de indivíduos diferentes, 57 indivíduos na base de dados de música e matemática e somente 17 indivíduos na base de dados de jogos de video game, para que houvesse um mesmo número de sinais de EEG de indivíduos para cada categoria de atividade (jogar video game; ouvir música; responder questões matemáticas), considerou-se 57 sinais coletados durante a atividade de jogar video game, pois os 17 participantes da base de dados foram submetidos a várias seções de coletagem de EEG, cada coleta em momentos diferentes, assim as três categorias de atividade tiveram a mesma quantidade de sinais de EEG, 57 sinais de cada atividade. Esse procedimento de balanceamento foi necessário para que os dados utilizados tivesse a mesma quantidade de sinais de EEG pra cada classe de atividade cognitiva, e assim o treino e teste do modelo não fosse prejudicado pela diferença de quantidade de voluntários de cada base de dados.

2.4 Etapa 4 - O modelo classificador

A partir do banco de dados criado, com todos os dados referentes às regiões anatômicas para os indivíduos das bases de dados, será possível desenvolver um modelo baseado em aprendizado de máquinas para classificar a atividade encefálica dentro de uma das três categorias: “Jogo de video game, Música, Matemática”. Nesta etapa, foram testados os seguintes modelos:

1. **Random Forest Classifier:** consiste em uma combinação de classificadores de árvore onde cada classificador é gerado usando um vetor aleatório amostrado independentemente do vetor de entrada [17]. Esse classificador aleatório usa o Índice de Gini como uma medida de seleção de atributo, que mede a impureza de um atributo em relação as classes. Para um determinado conjunto de treinamento T , selecionando um caso aleatoriamente e dizendo que ele pertence a alguma classe, o índice de Gini pode ser escrito como [18]:

$$\sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|)$$

onde $(f(C_i, T)/|T|)$ é a probabilidade de o caso selecionado pertencer à classe C_i .

2. **K-vizinho mais próximo (kNN):** A classificação K-vizinho mais próximo (kNN) é um dos métodos de classificação mais fundamentais e simples de classificação [19]. O classificador k-vizinho mais próximo é comumente baseado na distância euclidiana entre uma amostra de teste e as amostras de treinamento especificadas [20]. Seja x_i uma amostra de entrada com p recursos $(x_{i1}, x_{i2}, \dots, x_{ip})$, n seja o número total de amostras de entrada ($i = 1, 2, \dots, n$) e p o número total de recursos ($j = 1, 2, \dots, p$). A distância euclidiana

$d(x_i, x_l)$ entre a amostra x_i e $x_{l(l=1,2,\dots,n)}$ é definida como:

$$d(x_i, x_l) = \sqrt{((x_{i1}x_{l1})^2 + (x_{i2}x_{l2})^2 + \dots + (x_{ip}x_{lp})^2)}$$

3. **Linear Discriminant Analysis (LDA):** LDA é um algoritmo discriminante de subespaço clássico e representativo que tem sido amplamente utilizado em muitas aplicações [21]. Dado um conjunto de dados $X = [x_1, x_2, \dots, x_n] \in R_{mn}$ de classes C , LDA define uma matriz de dispersão dentro da classe S_w e uma matriz de dispersão entre as classes S_b . A transformação do espaço para a redução de dimensionalidade pode ser obtido resolvendo a seguinte função[21, 22]:

$$S_w^{-1} S_b Q = Q A$$

onde A é a matriz de autovalores de $S_w^{-1} S_b$ e Q é a matriz de transformação do objeto, que consiste em autovetores de $S_w^{-1} S_b$ (i.e., q_1, q_2, \dots, q_d) onde $d(d < m)$ possui valores próprios diferentes de zero.

4. **Decision Tree Classifier:** este método usa uma árvore de decisão (como um modelo preditivo) para ir de observações sobre um item (representado nos ramos) para conclusões sobre o valor alvo do item (representado nas folhas) [23]. As árvores de decisão estão entre os algoritmos de aprendizado de máquina mais populares devido à sua inteligibilidade e simplicidade[24].
5. **SVM - Support Vector Machines:** as máquinas de vetores de suporte (SVMs), introduzidas pela primeira vez por Vapnik [25], mostraram sua eficácia em muitos problemas de reconhecimento de padrões [26], e podem fornecer melhores desempenhos de classificação do que muitas outras técnicas de classificação. Um classificador SVM realiza a classificação binária, ou seja, ele separa um conjunto de vetores de treinamento para duas classes diferentes $(x_1, y_1), (x_2, y_2), \dots, (x_m, y_m)$, onde $x_i \in R_d$ denota vetores em um d -espaço de recurso dimensional e $y_i \in -1, +1$ é um rótulo de classe. O modelo SVM é gerado mapeando os vetores de entrada em um novo espaço de características de dimensão superior denotado como : $R_d \rightarrow H_f$ onde $d < f$. Então, um hiperplano de separação ótimo no novo espaço de recursos é construído por uma função kernel $K_{(x_i, x_j)}$, que é o produto dos vetores de entrada x_i e x_j e onde $K_{(x_i, x_j)} = (x_i) \cdot (x_j)$ [26].

A partir dos testes dos modelos, elegeu-se o modelo que teve melhor performance e para complementar e melhorar o classificador, usamos validação cruzada k -fold com 10 iterações. Este método é adequado, pois fornece informações sobre a robustez de um modelo. Neste tipo de validação, o conjunto de treinamento é dividido em 10 amostras escolhidas aleatoriamente. Nove delas são usadas para treinar o modelo, e o restante é usado para testá-lo. O resultado das 10 validações cruzadas é o valor médio de todos os ensaios.

A partir da validação cruzada obtemos uma matriz de confusão com os seguintes campos:

- Verdadeiro positivo: a atividade eletroencefálica pertence à categoria de atividade testada (Jogo, música ou matemática) e o modelo identificou corretamente.

- Falso negativo: a atividade eletroencefálica pertence à categoria de atividade, mas o modelo identificou erroneamente como pertencente à outra categoria de atividade.
- Falso positivo: a atividade eletroencefálica não pertence à categoria, mas o modelo identificou erroneamente como pertencente.
- Verdadeiro negativo: a atividade eletroencefálica não pertence à categoria e isso foi detectado corretamente pelo modelo.

O resultado dessa validação é um modelo válido, que pode ser usado para previsão no conjunto de teste real.

3 Resultados

Após todas as etapas para elaboração do modelo de classificação das atividades encefálicas, dentre as classes: “Jogo de Video Game, Música, Matemática”, obteve-se o modelo de classificação apresentado a seguir.

3.1 Classificador da atividade cognitiva a partir das regiões encefálicas ativadas

O uso da estimação de fontes encefálicas foi essencial na execução deste trabalho, pois a partir desta etapa foi possível identificar as regiões anatômicas que estão diretamente relacionadas com cada uma das atividades “Jogar video game, Música, Matemática”. A partir dessas regiões foi elaborado o banco de dados usado no treino e teste do modelo.

Na etapa de implementação do modelo, este foi implementado utilizando-se da biblioteca *pycaret* da linguagem *Python* [27], onde a partir do uso da biblioteca o modelo foi implementado, testado e validado com validação cruzada a partir do banco de dados produzido. O banco de dados foi dividido na porcentagem 90% para teste e 10% para treino, assim, a partir do banco de dados com as regiões encefálicas associadas a cada uma das atividades “Jogo de video game, Música, Matemática”, foi feito o treino e teste do modelo.

Considerando-se um modelo classificador multiclasse, fez-se um teste dos modelos de aprendizado de máquina que foram apresentados na seção 3. Estes modelos foram treinados e testados para os dados coletados e os resultados do treinamento estão mostrados na Tabela 1.

A partir do treinamento dos modelos, observou-se que o modelo que apresentou os melhores resultados na classificação do tipo de atividade executada a partir das regiões encefálicas relacionadas com as atividades “Jogo de video game, Música, Matemática” foi o modelo Random Forest, que foi o modelo escolhido.

Com o modelo Random Forest selecionado, pode-se observar a curva de aprendizagem do modelo, mostrada na Figura 2 onde observa-se um modelo com uma aprendizagem precisa e progressiva. Na Figura 2 podemos observar que a precisão do treino para o modelo está muito próxima de 100%, e que a validação do modelo tem um comportamento parecido.

Table 1: Resultados alcançados pelos modelos testados

| Modelo | Acurácia | Precisão |
|------------------------------|----------|----------|
| Random Forest Classifier | 0.9990 | 0.9933 |
| K Neighbors Classifier | 0.9900 | 0.9917 |
| Linear Discriminant Analysis | 0.9900 | 0.9917 |
| Decision Tree Classifier | 0.9800 | 0.9833 |
| SVM - Linear Kernel | 0.9800 | 0.9833 |



Fig. 2: Curva de aprendizagem do modelo Randon Forest

Após a validação do modelo, pode-se observar as regiões de maior influência na classificação do modelo de acordo com cada classe: jogo de video game, matemática e música. Na Figura 3 podemos observar as regiões com maiores influências na classificação para cada classe, na imagem, a classe 0 se trata da atividade de ouvir música, na classe 1 é a atividade de responder questões matemáticas e na classe 2 é a atividade de jogar video game.

As regiões encefálicas citadas na Figura 3 que apresentam relação positiva com a atividade se apresentam ativadas durante a execução da atividade específica, uma relação positiva significa que a ativação daquela região é um fator considerado na classificação do tipo de atividade. As regiões encefálicas que apresentam relação negativa com a atividade não apresentam atividade encefálica relacionada com a execução da atividade. Dentre todas as regiões que possuem relação positiva com a atividade, se destaca a região Inferior Parietal Esquerda, que é uma região do lóbulo parietal que se relaciona com emoção e coordenação motora [9, 13, 8, 6, 7]. Quatro das regiões com relação positiva com as atividades pertencem ao lóbulo frontal, ou seja, quatro regiões do lóbulo frontal apresentam alguma relação direta de ativação durante o desenvolvimento das atividades cognitivas: jogo de video game, música e matemática. Essas regiões do lóbulo frontal são relacionadas, segundo a literatura, com a cognição e aprendizagem [5]. Regiões do lóbulo occipital também se relacionam com as atividades de forma

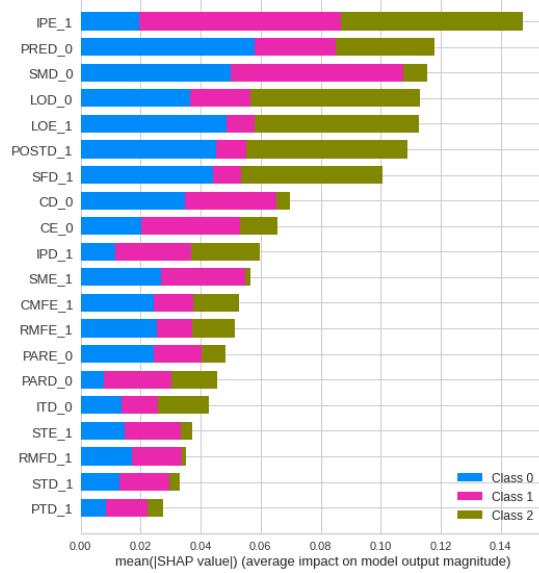


Fig. 3: Regiões encefálicas com maior influência na classificação da atividade realizada. IPE - Inferior Parietal Esquerdo; PRED - Precuneus Direito; SMD - Supra Marginal Direito; LOD - Lateral Occipital Direito; LOE - Lateral Occipital Esquerdo; POSTD - Postcentral Direito; SFD - Superior Frontal Direito; CD - Cuneus Direito; CE - Cuneus Esquerdo; IPD - Inferior Parietal Direito; SME - Superior Marginal Esquerdo; CMFE - Caudal Médio Frontal Esquerdo; RMFE - Rostral Médio Frontal Esquerdo; PARE - Parahippocampal Esquerdo; PARD - Parahippocampal Direito; ITD - Inferior Temporal Direito; STE - Superior Temporal Esquerdo; RMFD - Rostral Médio Frontal Direito; STD - Superior Temporal Direito; PTD - Posterior Temporal Direito. As regiões que apresentam uma relação positiva com a atividade em questão se apresenta seguida do valor 1 no eixo y, enquanto as regiões com relação negativa se apresentam com o valor 0 no eixo y.

positiva, o que indica uma ativação de regiões relacionadas a processamento de imagens [28].

Title Suppressed Due to Excessive Length

A partir da indicação de relação entre região encefálica ativada e atividade cognitiva determinada utilizando o modelo de classificação eleito: Random Forest, pode-se analisar a eficiência do modelo na classificação das atividades cognitivas a partir da ativação encefálica durante a realização das atividades classificadas por meio da curva ROC do método. A curva ROC do modelo, na Figura 4, indica a precisão de classificação do modelo, apresentando uma excelente classificação multiclasse.

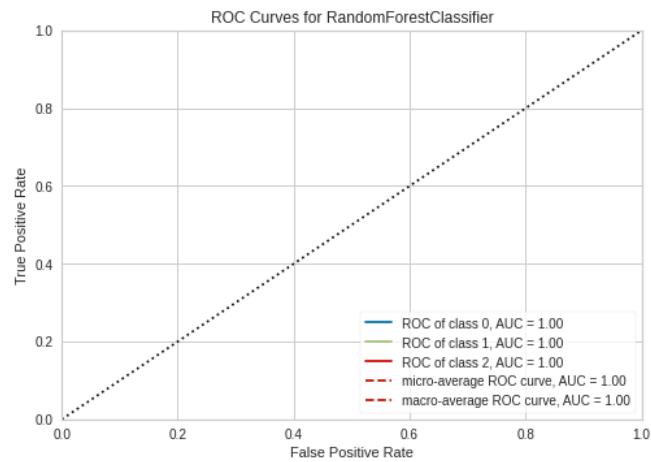


Fig. 4: Curva ROC do modelo.

Após a etapa de treino e teste do modelo, foi feito a observação da clusterização do modelo, que indica a forma como o modelo está dividindo as classes, ou como está a classificação do modelo quanto a cada uma das atividades “Jogo de video game, Música e Matemática”, e a Figura 5 mostra como está sendo a classificação do modelo. A clusterização consiste na apresentação da distribuição das atividades desempenhadas, mostrando a dispersão das atividades.

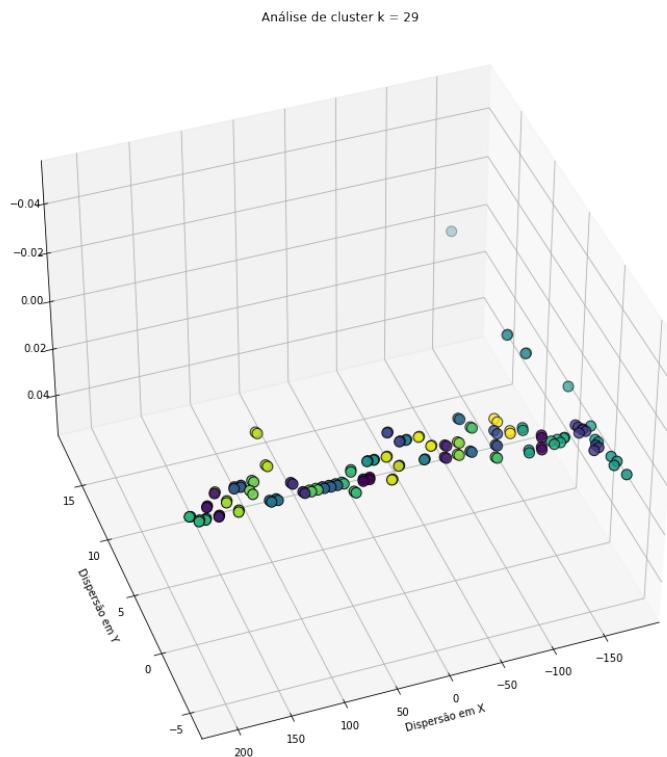


Fig. 5: Teste do modelo Randon Forest - realizado por meio da clusterização. A cor azul representa a classe de Jogar video game, a cor amarela a classe de ouvir música e a cor verde a classe de responder questões de matemática.

3.2 Discussão dos Resultados

Neste trabalho foi apresentado o desenvolvimento de um modelo de classificação de atividades encefálicas, baseado em um modelo de aprendizado de máquinas para classificar a atividade encefálica como pertencente a uma das três classes: Jogo de video game, música ou matemática. No modelo, considera-se o desenvolvimento de uma das três atividades pelo indivíduo. O modelo trabalha a partir do EEG coletado, com o uso da técnica de estimativa de fontes encefálicas, localizando a regiões encefálicas relacionadas a atividade.

O modelo de classificação das atividades encefálicas se mostrou eficiente, com um número de acertos de 0,9990 na fase de testes, e com um funcionamento adequado. Notamos que é possível a predição do tipo de atividade cognitiva que está sendo realizada a partir das regiões encefálicas ativadas que possuem relação com a tarefa em questão. Assim, este trabalho representa um passo importante no mapeamento das atividades cognitivas e sua relação com a atividade elétromagnética encefálica.

Ressaltamos, que há diversos trabalhos na literatura que relacionam a ativação encefálica com a atividade desempenhada [6, 7, 9], mas, em sua grande maioria, esses trabalhos utilizam imagem de Ressonância Magnética Funcional (fMRI), que é uma técnica cara, cuja manutenção e manuseio requer alto grau de conhecimento e habilidade. Neste trabalho, propomos o uso dos sinais de EEG, nos casos em que seja possível a substituição da fMRI, como fonte de observação da ativação encefálica, notando que este tipo de observação através dos sinais de EEG é possível com o uso da técnica de estimativa de fontes encefálicas. As vantagens do uso do sinais de EEG neste tipo de análise está relacionada com seu baixo custo, pouca manutenção, e facilidade de acesso, uma vez que esta é uma técnica bastante utilizada em todas as classes sociais para aquisição de informações quanto a saúde encefálica.

Observamos que o modelo classificador teve um desempenho bom e pretende-se, em trabalhos futuros, buscar o desenvolvimento de uma interface gráfica que possa classificar o tipo de atividade cognitiva em tempo real a partir do EEG captado, o que se mostra um grande diferencial no sentido de um monitoramento eficiente e rápido da atividade elétromagnética em tempo real, utilizando o modelo classificador aqui desenvolvido. Este é um passo importante na busca por um monitoramento seguro de pacientes que possuam alguma patologia que afete seu desenvolvimento cognitivo, uma vez que a atividade encefálica elétromagnética em indivíduos com problemas cognitivos não ocorre da maneira normal esperada, como em indivíduos normais.

4 Conclusões

Este trabalho apresentou o desenvolvimento de um modelo de classificação de atividades cognitivas a partir das regiões encefálicas relacionadas a atividade de três classes (Jogo de video game, música, matemática). Foi realizado o teste de alguns modelos de aprendizado de máquinas, mas o modelo eleito que apresentou o melhor desempenho foi o Random Forest utilizado no desenvolvimento

do modelo classificador. O modelo desenvolvido classifica a atividade encefálica como pertencente a uma das classes: jogo de video game, música, matemática.

Assim, este estudo representa um passo importante em mapear a atividade neural do funcionamento do encéfalo no contexto do desenvolvimento cognitivo dentro dos indivíduos, uma vez que ainda é um mistério como se comporta o cérebro de diferentes indivíduos no desenvolvimento cognitivo. A partir desse protótipo inicial, futuramente, pretende-se desenvolver uma interface que possa realizar a análise em tempo real utilizando o modelo de classificação aqui desenvolvido.

References

1. GIRI, Amita; KUMAR, Lalan; GANDHI, Tapan Kumar. Brain Source Localization in Head Harmonics Domain. *IEEE Transactions on Instrumentation and Measurement*, v. 70, p. 1-10, 2020.
2. MORIDERA, Takayoshi *et al.* High-resolution EEG source localization in segmentation-free head models based on finite-difference method and matching pursuit algorithm. *Frontiers in Neuroscience*, v. 15, p. 789, 2021.
3. Kurkin, Semen, *et al.* "Localizing oscillatory sources in a brain by meg data during cognitive activity." *2020 4th International Conference on Computational Intelligence and Networks (CINE)*. IEEE, 2020.
4. Cavanagh, James F., and Joel Castellanos. "Identification of canonical neural events during continuous gameplay of an 8-bit style video game." *Neuroimage* 133 (2016): 1-13.
5. MCCRAE, Christina S. *et al.* Neural activation changes in response to pain following cognitive behavioral therapy for patients with comorbid fibromyalgia and insomnia: a pilot study. *Journal of Clinical Sleep Medicine*, p. jscm. 9540, 2021.
6. ZHANG, Niegang *et al.* The effects of age on brain cortical activation and functional connectivity during video game-based finger-to-thumb opposition movement: A functional near-infrared spectroscopy study. *Neuroscience Letters*, v. 746, p. 135668, 2021.
7. ZHANG, Jin-Tao *et al.* Effects of craving behavioral intervention on neural substrates of cue-induced craving in Internet gaming disorder. *NeuroImage: Clinical*, v. 12, p. 591-599, 2016.
8. VOGEL, Stephan E.; DE SMEDT, Bert. Developmental brain dynamics of numerical and arithmetic abilities. *npj Science of Learning* v. 6, n. 1, p. 1-11, 2021.
9. LIU, Ying *et al.* Neural activation of different music styles during emotion-evoking. *Psychology of Music*, p. 0305735620968260, 2020.
10. Tadel, François, *et al.* "Brainstorm: a user-friendly application for MEG/EEG analysis." *Computational intelligence and neuroscience* (2011).
11. Yulin Wang and Wei Duan and Lihong Ding and Debo Dong and Xu Lei (2021). A test-retest resting and cognitive state EEG dataset. *OpenNeuro*. [Dataset] doi: 10.18112/openneuro.ds003685.v1.0.2
12. DUAN, Wei *et al.* Reproducibility of power spectrum, functional connectivity and network construction in resting-state EEG. *Journal of Neuroscience Methods*, v. 348, p. 108985, 2021.
13. WILKEY, Eric D. *et al.* The effect of visual parameters on neural activation during nonsymbolic number comparison and its relation to math competency. *NeuroImage*, v. 159, p. 430-442, 2017.

Title Suppressed Due to Excessive Length

14. James F Cavanagh and Joel Castellanos (2021). EEG: Continuous game-play of an 8-bit style video game. *OpenNeuro*. [Dataset] doi: 10.18112/openneuro.ds003517.v1.1.0
15. JATOI, Munsif Ali *et al.* EEG based brain source localization comparison of sLORETA and eLORETA. *Australasian physical engineering sciences in medicine*, v. 37, n. 4, p. 713-721, 2014.
16. Klein, Arno, *et al.* "Mindboggling morphometry of human brains." *PLoS computational biology* 13.2 (2017): e1005350.
17. BREIMAN, Leo. Random forests. *UzC Berkeley TR567*, 1999.
18. Kumar, Jothi Letchumy Mahendra, *et al.* "The classification of EEG-based winking signals: a transfer learning and random forest pipeline." *PeerJ* 9 (2021): e11182.
19. CHOUBEY, Hemant; PANDEY, Alpana. A combination of statistical parameters for the detection of epilepsy and EEG classification using ANN and KNN classifier. *Signal, Image and Video Processing*, v. 15, n. 3, p. 475-483, 2021.
20. RASHEED, Muneera A. *et al.* Use of artificial intelligence on Electroencephalogram (EEG) waveforms to predict failure in early school grades in children from a rural cohort in Pakistan. *Plos one*, v. 16, n. 2, p. e0246236, 2021.
21. MA, Delu *et al.* The Automatic Detection of Seizure Based on Tensor Distance And Bayesian Linear Discriminant Analysis. *International Journal of Neural Systems*, p. 2150006-2150006, 2021.
22. ZOLLANVARI, Amin; ABIBULLAEV, Berdakh. Bias Correction for Linear Discriminant Analysis. *Pattern Recognition Letters*, 2021.
23. SAFAVIAN, S. Rasoul; LANDGREBE, David. A survey of decision tree classifier methodology. *IEEE transactions on systems, man, and cybernetics*, v. 21, n. 3, p. 660-674, 1991.
24. FRAIWAN, Luay; HASSANIN, Omnia. Computer-aided identification of degenerative neuromuscular diseases based on gait dynamics and ensemble decision tree classifiers. *Plos one*, v. 16, n. 6, p. e0252380, 2021.
25. Vapnik V. *Statistical Learning Theory*. John Wiley; 1998.
26. NARAYAN, Yogendra. Motor-Imagery EEG Signals Classificationusing SVM, MLP and LDA Classifiers. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, v. 12, n. 2, p. 3339-3344, 2021.
27. Python Software Foundation. Referência da linguagem Python, versão 3.8.8 Disponível em <http://www.python.org>.
28. EVANS, S. L. *et al.* Evening preference correlates with regional brain volumes in the anterior occipital lobe. *Chronobiology International*, v. 38, n. 8, p. 1135-1142, 2021.

A Random LSTM Model for Stock Market Prediction

Elaine Pinto Portela¹ and Omar Andres Carmona Cortes²[0000–0002–5805–2490]

¹ Programa de Pós-Grauação em Engenharia da Computação e Sistemas (PECS), Universidade Estadual do Maranhão (UEMA), Cidade Universitária Paulo VI, São Luis, São Cristovão, MA, Brazil
elainepportela@gmail.com

² Departamento de Computação (DComp), Instituto Federal do Maranhão (IFMA), Av. Getulio Vargas, 04, Monte Castelo, São Luis, MA, Brazil
omar@ifma.edu.br

Abstract. This work proposes the Random LSTM model, a bagging-based ensemble algorithm, for predicting the price of BVSP (Ibovespa) and S&P500 stocks. Similar to Random Forest, in Random LSTM, we create random LSTM neural networks with only one hidden layer. Afterward, we train them using different samples from the training dataset in the ratio of 2/3. Then the final prediction is made by averaging the results of each LSTM. Results have shown that RLSTM presented better RMSE and MAE in both *BVSP* and *S&P500*.

Keywords: Deep Learning · Random LSTM · Time Series · Stock Prediction.

1 Introduction

In summary, a time series is a sequence of data in which we try to use the past to predict the future to perform decision-making properly. In other words, the idea is to extract unknown patterns and information from them, which is the central core of many real-world applications [20]. Figure 1 shows an example of time series depicting the evolution of the Amazon stock (AMZN) price from January-2015 to December-2019. In this example, we could use part of this time series to predict the price in the future. In fact, after creating a model, predictions can be made even in a future that did not happen yet, which would be the task in a real-world application.

As we can see in the previous example, an application can be built to predict the time series of stock market data. Doing so presents some exciting challenges. The primal one comes from the dynamics of the boisterous stock market. In other words, some facts that cannot be predicted can happen. For instance, if a CEO is caught committing fraud, which is an unpredictable fact, the value of the asset of this company will unfortunately dropdown.

The second one appears from the fact that time series usually contain sequential dependencies among the attributes, i.e., the sequence of each attribute



Fig. 1. Amazon stock (AMZN) price in dollars form January-2015 to December-2019

cannot be changed. For example, the values on successive time-stamps are closely related to one another [1]. Thus, if we use it as independent features, the relationship between these values can be lost. Consequently, this kind of prediction is one of the most difficult in financial applications.

Despite the difficulties of modeling the stock market, dealing with assets (sell and buy) is essential to economic growth because it grants the opportunity to investors that provide the necessary resources for the growth of companies [17]. In other words, it is a two-way road because the company can attract some investments while clients expect to receive future profit.

Therefore, considering that the stock market is composed of time-series data, we can use several algorithms to deal with it. A common way of doing so is using recurrent neural networks. However, recurrent multilayer neural networks, as well as multilayer perceptron (MLP), suffer from the vanishing gradient problem [16], which can be addressed by using the so-called Long Short-Term Memory (LSTM) neural networks, a variation of recurrent neural network.

In this context, this work proposed a new algorithm called Random LSTM, or RLSTM, a bagging ensemble machine learning model similar to Random Forest. The essential idea is to instantiate several independent models. The prediction is then made by votation, in the case of classification problems, or by averaging the results in the case of regression problems.

Thus, this work is divided as follows: Section 2 presents the related works and highlights the contributions of this investigation. Section 3 shows how the LSTM works. Section 4 presents the configuration of the experiment, the used dataset, and the results. Finally, in Section 5 we discuss the results, and then we present the conclusions of this work.

2 Related Works

Machine learning for predicting the stock market has received significant attention using algorithms such as Decision Trees (DT) [12], Support Vector Machine

A Random LSTM Model for Stock Market Prediction

(SVM) [7], and Neural Networks (NN) [21]. Further, Nti et al. [15] claim in their investigation that SVM and NN are the principal machine learning algorithms used for this particular task. On the other hand, deep learning has emerged as the best performing predictor within the machine learning field, especially in the finance area. In fact, Sezer et al. [19] have proven that deep learning models perform better than traditional machine learning algorithms.

Thus, considering that time series are sensitive to historical data, and also taking into account that ANNs can suffer from vanishing gradients. The natural path is to use recurrent neural networks, especially those that solve the vanishing gradient problem, the LSTM.

The LSTM networks was introduced by Hochreiter & Schmidhuber [6] in 1997. Since then, many investigations have been conducted using this kind of neural network in many different areas. For example, Ma et al. [11] uses an LSTM for predicting traffic speed; Cai & Liu used LSTM in speech recognition; Kim & Cho [9] proposed an LSTM combined with a Convolutional Neural Network (CNN) for prediction residential energy consumption; and, Jorges et al. [8] used an LSTM for prediction and reconstruction of ocean wave heights.

Regarding stock market pricing, Nelson et al. [14] used LSTM to predict the price of some IBOVESPA assets, such as, BOVA11, BBDC4, CIEL3, ITUB4 and PETR4, showing better accuracy than Random Forest and Multilayer Perceptron (MLP). Baek & Kim [3] proposed a framework for making prediction in S&P500 and KOSPI200 datasets using LSTM models. Results have shown that the proposed framework presented better results than a regular Deep Neural Network (DNN) and a RNN. Setia & Raut [18] showed that LSTM and GRU (another type of RNN) presented the best results when forecasting prices in S&P500 if compared against SVM and a DNN with four layers and 64, 128, and 256 neurons on each one, respectively. Portela & Cortes [17] used an LSTM composed by eight layers for predicting the price of all stock in S&P500.

To the best of our knowledge, the only works that investigates similar approaches are Al Rafi's et al. [2] and Koteeti's et al.[10]. However, their approaches are much more complex. In the first one, the authors propose an algorithm called Random Weighted LSTM, in which the dataset is divided into two subsets then the algorithm tries to find out the best subset samples for training weights and the model. Then the best subsets passed through an LSTM. Finally, the weighted model performs a classification task. Hence, as we can see, this approach goes differently, paying attention to the random subsets and not to random LSTM models.

In the second approach, Koteeti's et al. propose a bagging ensemble algorithm devised by different neural networks: RNN, GRU, LSTM, and Bi-RNN. In fact, the proposal is a bagging algorithm with a stacking appearance because they used independent and different machine learning algorithms, which is essentially different than we are proposing.

Thus, unlike those previous investigations that use LSTM directly and differently from Rafi's work, which is a classification algorithm and essentially different, we proposed an ensemble bagging-based algorithm in which many simple

LSTM models containing only one hidden layer are randomly created. Then, each model is trained using different sets of samples from the database. The regression is done by averaging the output of each model as happens in the random forest algorithm.

3 The Random LSTM

3.1 LSTM

An LSTM network is a recurrent neural network capable of dealing with numerical or symbolic inputs to predict future data. Unlike traditional recurrent neural networks, LSTM can store long-time memory, being useful for prediction using long-time data [13]. In other words, LSTMs are suitable to scenarios in which a significant gap exists between the current information and the point where it is required.

Figure 2 illustrates an LSTM architecture, in which we can see four gates: forget gate, input gate, control gate, and output gate. Moreover, we can see the structures vector multiplication, vector addition, sigmoid and tanh neural network.

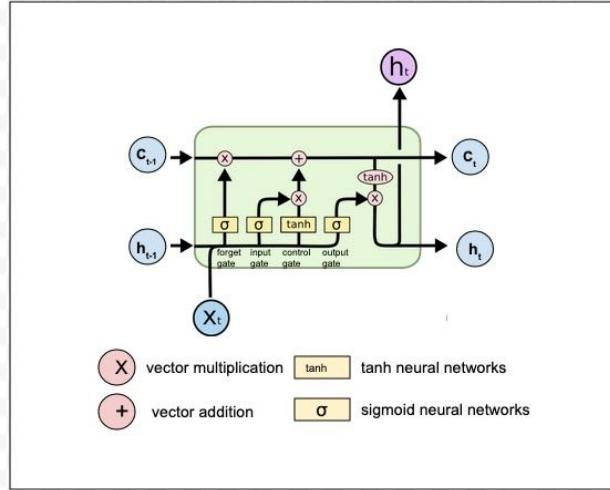


Fig. 2. LSTM Architecture

Three gates use the sigmoid activation function (σ) for processing the information. Thus, the output is either zero or one. Zero means that the gate is blocking the information. On the other hand, one indicates that the information is passing through.

$$\begin{aligned} f_t &= \sigma(W_f [h_{t-1}, x_t] + b_f) \\ i_t &= \sigma(W_i [h_{t-1}, x_t] + b_i) \\ o_t &= \sigma(W_o [h_{t-1}, x_t] + b_o) \end{aligned} \quad (1)$$

The three *sigma* outputs denoted by f_t , i_t , and o_t are expressed by Equations 1, in which w_k is the weight in the respective gate, h_{t-1} is the previous output at time $t-1$, x_t is the current input at time t , and b_k is the bias at the respective gate. The final output is depicted by Equations 2, in which C_t and C_{t-1} represent the new and previous cell state, respectively, \tilde{C}_t is a candidate to cell state at time t , and \times depicts a vector element-wise multiplication.

$$\begin{aligned} h_t &= o_t \times \tanh(C_t) \\ C_t &= f_t \times C_{t-1} + i_t \times \tilde{C}_t \\ \tilde{C}_t &= \tanh(w_c[h_{t-1}, x_t] + b_c) \end{aligned} \quad (2)$$

3.2 Our Proposal: Random LSTM

Ensemble algorithms aggregate a collection of learning models. There are mainly three types of ensembling. Stacking in which different and independent models are trained, then each model's results are combined for the final result. The second and third combination forms are bagging and boosting. In both, a set of simple models are created. In bagging (an acronym for bootstrap aggregating), all models are independent, and the task of classifying or regressing is done by combining the model's outputs. The combination depends on the task. In classification models, the combination is usually done by voting, in which the majority decides to which class the input belongs. In regression tasks, the final decision is performed by computing the average of each output. Finally, in boosting models, the output of a model is used as input of another one, then the final model performs the final classification or regression.

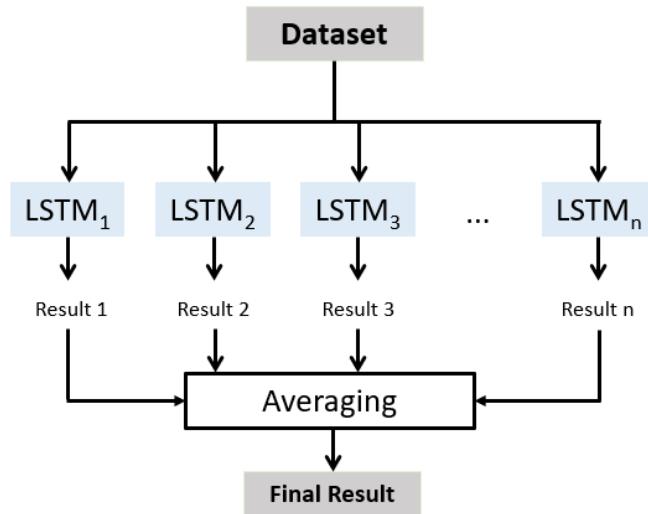
One of the most knowing ensemble algorithms is the Random Forest [4], a collection of random decision trees. In other words, instead of seeking to optimize a predictor as in traditional decision trees, the algorithm sets a pool of random individual trees, reaching better predictive performance in practice [5].

The Random LSTM algorithm is a bagging ensemble model. The idea is to create several LSTM randomly, *i.e.*, with a random architecture. Because bagging ensemble models produce small models, the Random LSTM creates simple neural networks composed of no more than two hidden layers. In other words, the proposed model is composed of many simple LSTM models with one or two layers. Moreover, each LSTM model has other components that can be chosen randomly, such as the activation function of each layer and the dropout rate. Table 1 summarizes the random elements and their possible values (domain).

Table 1. Random elements in Random LSTM

| Element | Domain |
|---------------------|--------------------------|
| Layers | {1,2} |
| Activation Function | {tanh, sigmoid} |
| Dropout | {0.10, 0.15, 0.20, 0.25} |

Moreover, similarly to Random Forest, each model is trained with different samples of the data set composed of 2/3 of the dataset. As each model provides a result, the final result is obtained by averaging all results as presented in Figure 3.

**Fig. 3.** Random LSTM Algorithm

4 Computational Experiments

This section presents the dataset that we used in the experiments (S&P500 and Ibovespa - BVSP), the metrics to evaluate the algorithm's performance, how the experiment was configurated, and the obtained results.

4.1 Dataset

Ibovespa (BVSP) index and S&P500 index price (SPX) compose the dataset from 2012/01/01 to 2020/27/01 as we can see in Figure 4. Both datasets can

A Random LSTM Model for Stock Market Prediction

be downloaded on WSJ website or be imported by *DataReader* library from Python.

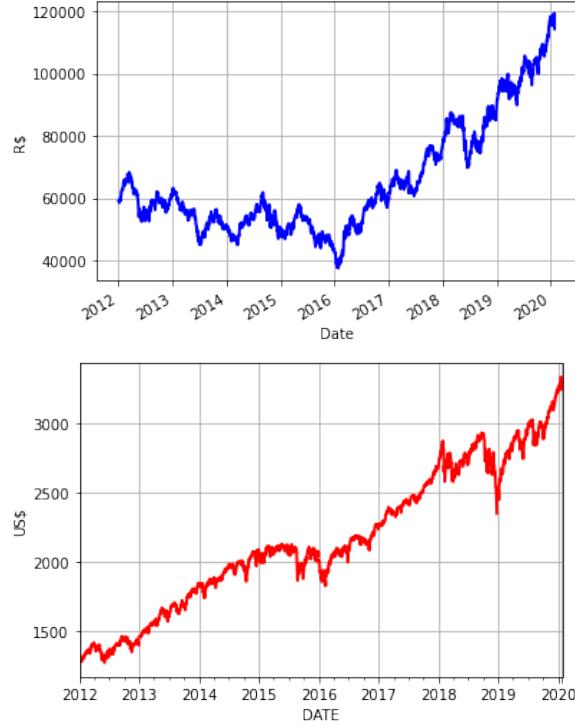


Fig. 4. BVSP and S&P500 price in dollars and reais, respectively form January-2012 to January-2020

The data holds the daily information of the open, high, low, and close of each index, as shown in Table 2. Moreover, as preprocessing, we removed the timestamp attribute. However, we are still obeying the data order. Furthermore, we added three new attributes as follows: mean, standard deviation, and variance of the four values of our database, open, high, low, and close. Finally, the stock values and the new attributes were normalized by using the standard deviation, *i.e.*, subtracting the mean (μ) from data (x) and then dividing it by the standard deviation (σ) as shown in Equation 3.

$$x = \frac{(x - \mu)}{\sigma^2} \quad (3)$$

Table 2. S&P500 stock data

| <i>date</i> | <i>open</i> | <i>high</i> | <i>low</i> | <i>close</i> |
|-------------|-------------|-------------|------------|--------------|
| 01/31/20 | 3282.33 | 3282.33 | 3214.68 | 3225.52 |
| 01/30/20 | 3256.45 | 3285.91 | 3242.80 | 3283.66 |
| 01/29/20 | 3289.46 | 3293.47 | 3271.89 | 3273.40 |
| . | . | . | . | . |
| 01/05/12 | 1277.30 | 1283.05 | 1265.26 | 1281.06 |
| 01/04/12 | 1277.03 | 1278.73 | 1268.10 | 1277.30 |
| 01/03/12 | 1258.86 | 1284.62 | 1258.86 | 1277.06 |

4.2 Metrics

The main metrics for evaluating regressive models are R squared (R^2), RMSE, and MAE. The R^2 metric assesses how well the model fits data, *i.e.*, how independent attributes explain the dependant variable. Its value is expressed in Equation 4 and its value is within the interval [0, 1], in which SSR is the total sum of square residual ($\sum(\hat{y}_i - \bar{y})^2$) and SST is the sum of square total ($y_i - \bar{y}^2$). Thus, the closer to one the R^2 , the better the model fits.

$$R^2 = \frac{\sum SSR}{\sum SST} \quad (4)$$

The next metric is the RMSE (Root Mean Squared Error) depicted in Equation 5, the square root of the mean of the squared errors. This metric indicates how close the predicted values are to the real ones. In this context, the closer to zero the RMSE, the better the results.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (5)$$

Finally, MAE is the mean of the absolute value of errors, *i.e.*, the difference between the actual and prediction as expressed in Equation 6. Thus, the prediction is better as MAE goes to zero.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

In Equations 4, 5, and 6, y_i represents the i^{th} value of the dependent variable, \bar{y} is the mean of the dependent variable, \hat{y} is the predicted variable, and \hat{y}_i is the i^{th} value of predicted dependent variable.

4.3 Setup

The experiments of this study were written using Python version 3.7.2 and Keras with Tensorflow version 2.6.0. The code ran in Google Colab Pro using the

A Random LSTM Model for Stock Market Prediction

available configuration of a virtual computer with 1-core Intel Xeon 2.3GHz, 12Gb of RAM, 167Gb of HDD, and a 16Gb GPU.

The configuration of the RLSTM models uses 10, 50, 100, 200, and 500 random models. The dataset from 2012 to 2016 was used as a training set. Then the predictions are made from 2017 to 2020, and the three metrics presented previously have been used. Moreover, the results are compared against a traditional LSTM with eight hidden layers, chosen by running some experiments. The training step was performed using 15 epochs, the Adam optimizer, using the close value as the value to be predicted.

4.4 Results

BVSP

Table 3 shows the considered metrics of the BVSP using a deep learning LSTM with 8 layers as proposed by Portela [17] and the RLSTM using 10, 50, 100, 200, and 500 random models, respectively. As we can see, the best R^2 is presented by the Deep Learning model. On the other hand, the best MAE and RMSE are shown by 100 random LSTM. In other words, the error is smaller in RLSTM.

Table 3. Metrics - Ibovespa BVSP

| Type | R^2 | MAE | RMSE |
|-----------|---------------|---------------|---------------|
| Deep LSTM | 0.9960 | 0.2971 | 0.2971 |
| 10 RLSTM | 0.9951 | 0.2581 | 0.2581 |
| 50 RLSTM | 0.9946 | 0.2543 | 0.2543 |
| 100 RLSTM | 0.9948 | 0.2525 | 0.2525 |
| 200 RLSTM | 0.9949 | 0.2549 | 0.2549 |
| 500 RLSTM | 0.9947 | 0.2536 | 0.2536 |

Figure 5 presents the prediction of each configuration in the period from 2017 to 2020. Visually, the results are pretty similar. However, MAE and RMSE indicate that RLSTM presents a smaller error, which can represent a meaningful amount of money if the financial investment is high.

In order to identify the difference that leads to a better RMSE and MAE in RLSTM, we plotted a delta graphic for BVSP in Figure 6. However, it is hard to see the difference because while RLSTM predicts above the close price, the Deep model predicts below.

S&P500

Table 4 illustrates the considered metrics of S&P500 using a deep learning LSTM with 8 layers and the RLSTM using 50, 100, 200, and 500 random models, respectively. As we can notice, the deep neural network presented the best R^2 . On the other hand, the LSTM using 10 random models shows better MAE and

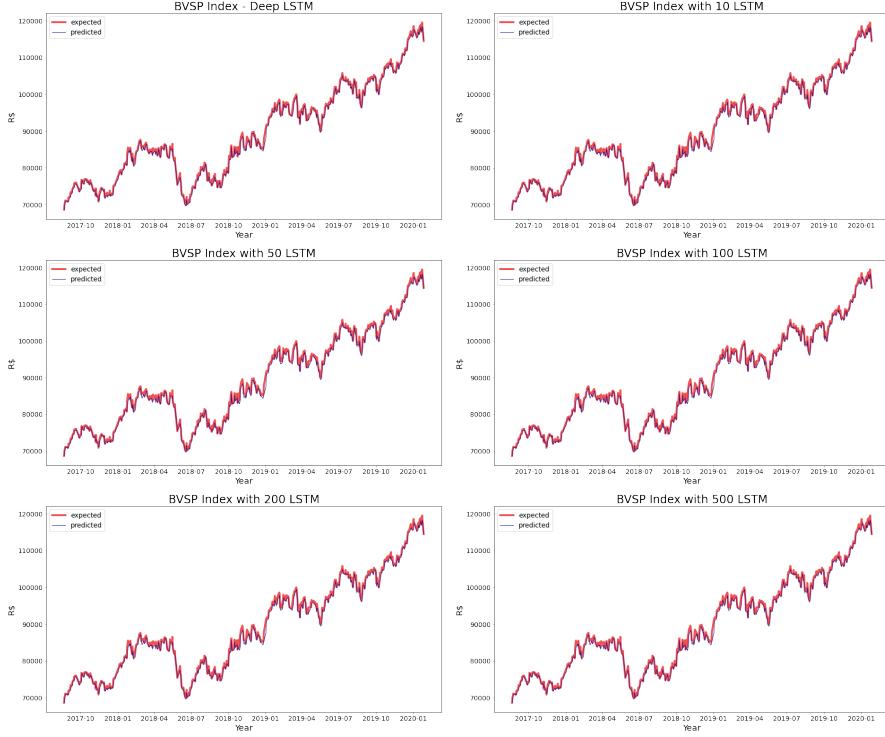


Fig. 5. Values from BVSP index prediction using a Deep LSTM, a Random LSTM with 10, 50, 100, 200 and 500, respectively.

RMSE. Thus, a shorter error can represent better predictions or returns if the amount of investments is high.

Figure 8 shows the prediction of the BVSP using a deep learning LSTM with 8 layers and the RLSTM using 50, 100, 200, and 500 random models, respectively. Similar to the previous case, the results are visually similar; however, as previously stated, the small difference in the error can express a higher amount of money when the investment is higher.

Also, as done previously, in order to identify the difference that leads to a better RMSE and MAE in RLSTM, we plotted a delta graphic for S&P 500 in Figure 6. However, it is hard to see the difference because while RLSTM tends to predict above the close price, the Deep model predicts much below.

5 Conclusions

This paper presented a new algorithm called Random LSTM. The central idea was to create several simple LSTM models containing random elements: one or two hidden layers, the activation function, and the dropout rate. Then, each

A Random LSTM Model for Stock Market Prediction

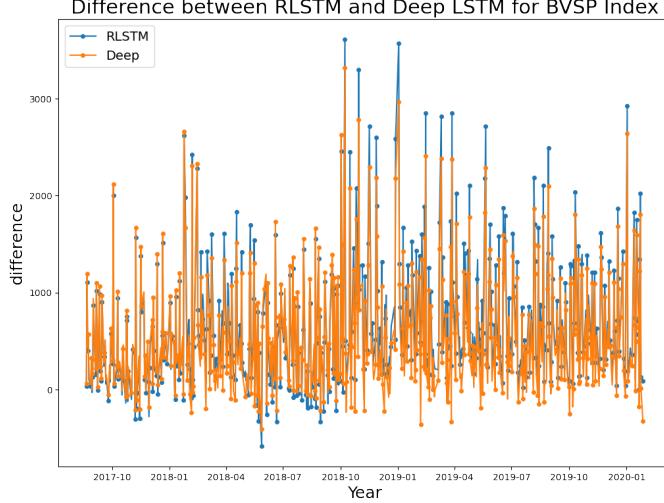


Fig. 6. Delta Plot with the 100 RLSTM and the Deep LSTM for the BVSP index.

Table 4. Metrics - S&P 500

| Type | R ² | MAE | RMSE |
|-----------|----------------|---------------|---------------|
| Deep LSTM | 0.9967 | 0.5441 | 0.5441 |
| 10 RLSTM | 0.9845 | 0.2550 | 0.2550 |
| 50 RLSTM | 0.9838 | 0.2769 | 0.2769 |
| 100 RLSTM | 0.9832 | 0.2772 | 0.2772 |
| 200 RLSTM | 0.9844 | 0.2840 | 0.2840 |
| 500 RLSTM | 0.9844 | 0.2807 | 0.2807 |

model is trained using a different sample from the database. Finally, the algorithms perform the regression task by averaging the output of each LSTM.

The models were evaluated using R^2 , RMSE, and MAE. In this context, it is essential to emphasize that the RMSE metric suits better in comparing two regression models. Furthermore, the MAE and RMSE measure the average and standard deviation of the absolute difference between the predicted and the expected, respectively, and the R Squared represents the proportion of the variance of the dependent variable.

The proposal was tested for stock market prediction using two indices, S&P500 and BVSP (Ibovespa). The BVSP index showed that the Random LSTM with 100 units presented a slighter error than a Deep LSTM. Indeed, all RLSTM presented a smaller error if compared against the traditional eight-layered LSTM model.

While in S&P500, the smallest error was presented by RLSTM using 10 units. Furthermore, it is impressive that in both *S&P500* and *BVSP*, the error (MAE and RMSE) is lesser than the half in RLSTM. Thus, all in all, even

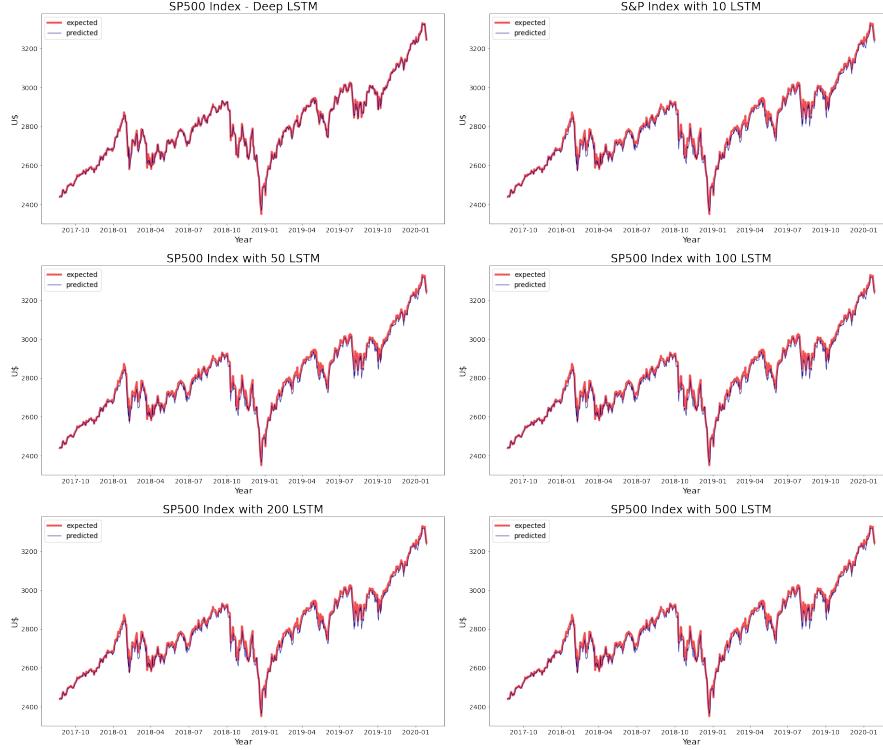


Fig. 7. Values from S&P500 index predicted using a Deep LSTM, a Random LSTM with 10, 50, 100, 200 and 500 respectively.

though the Deep Learning model with eight layers tended to present the best R^2 , RLSTM tended to present a smaller error which can represent a considerable amount of money if the investment in these assets is high. Nonetheless, additional investigation using more indices is necessary to evaluate the performance of the RLSTM.

Future work includes (i) evaluating the RLSTM using other indices; (ii) testing the prediction for optimizing a portfolio; and modifying the model in order to perform classification tasks; and (iii) adding more random features to LSTM, such as using different optimizers and different activation functions.

References

1. Aggarwal, C.C.: Neural Networks and Deep Learning: A Textbook. Springer (2018)
2. Al Rafi, A.S., Rahman, T., Al Abir, A.R., Rajib, T.A., Islam, M., Hossain Mukta, M.S.: A new classification technique: Random weighted lstm (rwl). In: 2020 IEEE Region 10 Symposium (TENSYMP). pp. 262–265 (2020)

A Random LSTM Model for Stock Market Prediction

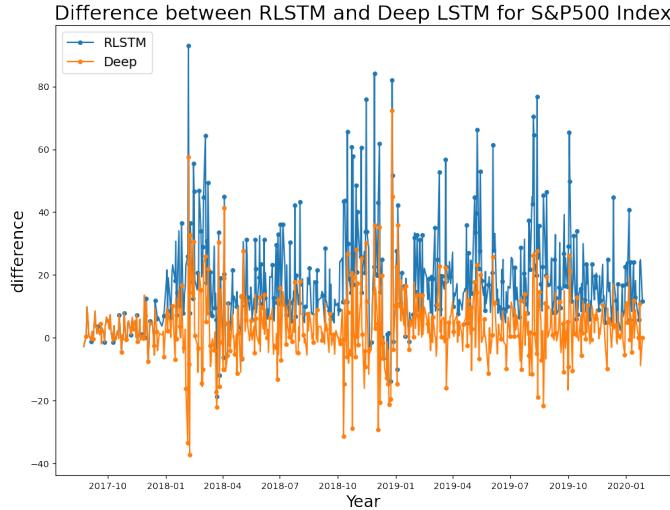


Fig. 8. Delta Plot with the 10 RLSTM and the Deep LSTM for the S&P500 index.

3. Baek, Y., Kim, H.Y.: Modaugnet: A new forecasting framework for stock market index value with an overfitting prevention lstm module and a prediction lstm module. *Expert Systems with Applications* **113**, 457–480 (2018)
4. Breiman, L.: Random forests. *Machine Learning* **45**(1), 5–32 (2001)
5. Genuer, R., Poggi, J.M.: *Random Forest With R*. Use R, Springer (2020)
6. Hochreiter, S., Schmidhuber, J.: Long Short-Term Memory. *Neural Computation* **9**(8), 1735–1780 (1997)
7. Hu, Z., Zhu, J., Tse, K.: Stocks market prediction using support vector machine. In: 2013 6th International Conference on Information Management, Innovation Management and Industrial Engineering. vol. 2, pp. 115–118 (2013)
8. Jörges, C., Berkenbrink, C., Stumpe, B.: Prediction and reconstruction of ocean wave heights based on bathymetric data using lstm neural networks. *Ocean Engineering* **232**, 109046 (2021). <https://doi.org/https://doi.org/10.1016/j.oceaneng.2021.109046>, <https://www.sciencedirect.com/science/article/pii/S0029801821004819>
9. Kim, T.Y., Cho, S.B.: Predicting residential energy consumption using cnn-lstm neural networks. *Energy* **182**, 72–81 (2019). <https://doi.org/https://doi.org/10.1016/j.energy.2019.05.230>, <https://www.sciencedirect.com/science/article/pii/S0360544219311223>
10. Kotteti, C.M.M., Dong, X., Qian, L.: Ensemble deep learning on time-series representation of tweets for rumor detection in social media. *Applied Sciences* **10**(21) (2020)
11. Ma, X., Tao, Z., Wang, Y., Yu, H., Wang, Y.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. *Transportation Research Part C: Emerging Technologies* **54**, 187–197 (2015)
12. Miró-Julià, M., Fiol-Roig, G., Isern-Deyà, A.P.: Decision trees in stock market analysis: Construction and validation. In: García-Pedrajas, N., Herrera, F., Fyfe,

- C., Benítez, J.M., Ali, M. (eds.) Trends in Applied Intelligent Systems. pp. 185–194. Springer Berlin Heidelberg, Berlin, Heidelberg (2010)
- 13. Moghar, A., Hamiche, M.: Stock market prediction using lstm recurrent neural network. *Procedia Computer Science* **170**, 1168–1173 (2020), the 11th International Conference on Ambient Systems, Networks and Technologies (ANT) / The 3rd International Conference on Emerging Data and Industry 4.0 (EDI40) / Affiliated Workshops
 - 14. Nelson, D.M.Q., Pereira, A.C.M., de Oliveira, R.A.: Stock market’s price movement prediction with lstm neural networks. In: International Joint Conference on Neural Networks (IJCNN). pp. 1419–1426 (2017)
 - 15. Nti, I., Adekoya, A., Weyori, B.: A systematic review of fundamental and technical analysis of stock market predictions. *Artificial Intelligence Review* **53**, 3007–3057 (2020)
 - 16. Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks. In: ICML Conference. pp. 1310—1318 (2013)
 - 17. Portela, E.P., Cortes, O.A.C.: Otimização de portfólio futuro baseado em aprendizagem profunda e algoritmo evolutivo multiobjetivo. *Revista Eletrônica de Iniciação Científica em Computação* **19**(1) (set 2021), <https://sol.sbc.org.br/journals/index.php/reic/article/view/2195>
 - 18. Sethia, A., Raut, P.: Application of lstm, gru and ica for stock price prediction. In: Satapathy, S.C., Joshi, A. (eds.) Information and Communication Technology for Intelligent Systems. pp. 479–487. Springer Singapore, Singapore (2019)
 - 19. Sezer, O.B., Gudelek, M.U., Özbayoglu, A.M.: Financial time series forecasting with deep learning : A systematic literature review: 2005-2019. CoRR **abs/1911.13288** (2019)
 - 20. Ye, Y., Jiang, J., Ge, B.: Similarity measures for time series data classification using grid representation and matrix distance. *Knowledge and Information Systems* **60**, 1105—1134 (09 2019)
 - 21. Yetis, Y., Kaplan, H., Jamshidi, M.: Stock market prediction by using artificial neural network. In: World Automation Congress (WAC). pp. 718–722 (2014)

COVID-19 Cognitive Sequelae and Their Possible Relation with Educational Issues: a systematic review

Drielle Viana Vieira¹, Tainá dos Santos Rêgo¹, Sayonara Pereira da Silva¹,
Maria Carolina Gonzalez¹, Edgard Morya¹,

¹ Edmond and Lily Safra International Institute of Neuroscience, Santos Dumont Institute. 1560 Alberto Santos Dumont Avenue, Macaíba, BR
drielle.vieira@edu.isd.org.br - <https://orcid.org/0000-0002-7295-2428>
taina.rego@edu.isd.org.br - <https://orcid.org/0000-0002-4161-8726>
sayonara.silva@edu.isd.org.br - <https://orcid.org/0000-0002-7865-1677>
carolina.gonzalez@isd.org.br - <https://orcid.org/0000-0003-4685-6739>
edgard.morya@isd.org.br - <https://orcid.org/0000-0003-0954-5317>

Abstract. The first case of COVID-19 was confirmed in China in December 2019. Since 2020, many studies around the world try to understand how SARS-CoV-2 infection acts on the nervous system, and how neuropsychological sequelae can vary from person to person. In this systematic review, we sought to correlate the occurrence of SARS-CoV-2 infection in/with memory deficits, and how this could lead to educational problems in the future. We included studies from June/2020 to October/2021 made in people who had the infection and recovered from it, reporting cognitive problems such as memory, difficulty concentrating, among other neuropsychological problems. The review was based on the state of the art of studies on the memory/COVID-19 relationship, and the following reports were analyzed: 1 case-control, 1 case report, and 12 cohort studies. At the end of this study, we propose that interactive games developed with Artificial Intelligence can contribute to ameliorate the deficits in concentration, learning and memory reported in patients recovered from COVID-19.

Keywords: Education, Memory, Covid-19, gamification.

1 Introduction

Origin of COVID

In December 2019, the first case of a respiratory syndrome caused by the new coronavirus was detected in China, in the city of Wuhan, which later became known as COVID-19. Rapidly spreading, the COVID-19 virus can be transmitted by inhalation or direct contact with infected droplets.[1]. The incubation period for SARS-CoV-2 ranges from 1 to 14 days [2]. Until the beginning of October 2021, 219.456.675 cases of COVID-19 were registered around the world, and 214.908.893 patients were reported as recovered with or without sequelae [3].

Post-COVID neuropsychological complaints

COVID-19 symptoms vary from fever and cough to myalgia and fatigue [3]. Among the discussions involving COVID-19, a lot is said about symptoms and syndromes acquired after COVID. People who had contact with the virus and that presented from mild to severe forms of the disease, describe different symptoms after SARS-CoV-2 infection, which are not restricted and limited to problems in the airways. Patients report lots of neurological issues after infection, such as headache, dizziness, altered mental status, fatigue, and memory problems. COVID-19 affects the human body in several ways. Among them, the one that most instigates researchers is how the infection can affect regions of the brain. [4].

Mechanism of viral influence on memory

Recent studies show how SARS-CoV-2 virus infection may be associated with neurological complications in the central and peripheral nervous systems [5]. This is because the virus can enter the human body through angiotensin-converting enzyme (ACE-2) receptors that are present in endothelial cells of brain vessels, with the possibility of the virus breaking and crossing the already damaged blood-brain barrier [6].

Viral infection is spread to the brain after a high viral load associated with the patient's immune vulnerability.

Baddeley et al. [7] considered that memory is a complex and multiple system combined by coding arrangements or subsystems that allow the storage and retrieval of information in the brain. Memory activation is attributed to several stimuli, being classified accordingly. Information is stored in different structures in the brain, but there are brain regions to which all the information converges.

These zones of convergence constitute the records of experiences, among these zones, the following brain areas of convergence can be mentioned: amygdala, orbitofrontal cortex, hippocampus, hypothalamus, among others [8]. Memory can be classified and divided as follows: explicit or implicit memory, or working memory, short-term and long-term memory, these last three refer to the time in which memories are stored [9].

Currently, it is known that neuroimmune cells are extremely necessary for normal memory function, related to the ability to retain knowledge [4]. When there is an infection in the body, all the signals the body gives are the result of protective responses by the human immune system. That is, specialized cells are activated when there is the detection of invasion by foreign bodies, sending large amounts of inflammatory signals, causing the way that neurons communicate to change. This sending of inflammatory signals can trigger a reaction where neuronal connections are destroyed [10].

Neuroimmune cells bind between neurons in response to disease and dump inflammatory signals at these junctions, causing disruptions in the connections between memory-storing neurons. These interruptions, with inflammatory signals, can cause permanent memory dysfunction due to damage to neural connections or neurons themselves, thus irreversible [4].

One of the main immune responses of the body to SARS-CoV-2

infection involves a large transmission of inflammatory signals sent by the brain, which can have the effect of a permanent long-term change associated with memory [4].

Nothing can be recognized or remembered without first being noticed [4]. Thus, we have that memory acquisition and learning go hand in hand, and it is concluded that one can only memorize what was once learned, lived, or experienced by human beings.

Learning is a consequence of storing information practiced, experienced, or experienced by human beings. The stored information that was generated by learning becomes what is called memory [9].

With the neurological sequelae left by the COVID-19 infection, reports of memory loss or difficulty in memorizing new information have been growing. However, it is not known for sure how many people were affected by this sequel around the world. We conducted this systematic review for trying to understand how COVID-19, being harmful to memory, can lead to future problems in education.

2 **Method**

The present systematic review examined studies with experimental or quasi-experimental designs that investigated cognitive aspects of COVID-19 survivors through the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [11].

Search strategy

The systematic search was carried out in October 2021, using three databases: PubMed, Scopus, and Web of Science. The following descriptors were used: "memory", "COVID-19", "sars-cov-2", "coronavirus", and "pandemic", filtered by title, abstract and keywords. Each base went through the same search phases, with the Boolean operators AND and OR. The survey was conducted in 5 steps: (1) memory AND COVID-19; (2) memory AND sars-cov-2; (3) memory AND coronavirus; (4) memory AND pandemics; (5) #1 OR #2 OR #3 OR #4.

The results obtained were attached to the RAYYAN QCRI tool [12] and evaluated by title and abstract based on the PICO strategy.

Inclusion criteria

Studies were incorporated: (a) performed with humans in any age group without a history of deficit or disorder in pre-covid cognitive functions; (b) that aimed to cognitively evaluate people recovered from COVID-19 infection; (c) in English; (d) published between June 2020 and October 2021.

Exclusion criteria

Surveys were conducted with: (a) patients hospitalized due to COVID-19;

(b) people infected with COVID-19 during the period the research was conducted; (c) editorials, reviews, systematic reviews, and protocols were disregarded.

Data selection

Two independent researchers, blinded to each other's decisions, separately assessed the found references. The blinding was broken by a third author, who resolved conflicts. Studies with at least two concordances were selected for a full reading. Methods Data were extracted from each publication, such as sample characteristics, information on cognitive assessment, evaluation period, main objective, and results concerning the cognitive assessment.

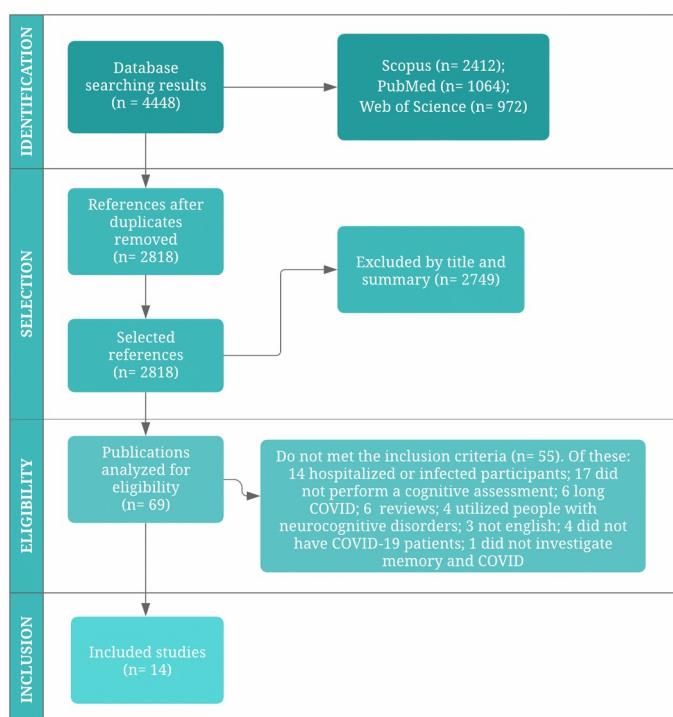
Risk bias

All of the included studies went to a blinded evaluation performed by two independent researchers. The references were evaluated through the versions for case-control, cohort, and cross-sectional of the Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses [13].

3 Results

The searches identified 4448 publications indexed in the Scopus ($n = 2412$), PubMed ($n = 1064$) and Web of Science ($n = 972$) databases. After initial screening, 69 studies were considered for analysis, from which 14 studies were extracted for meeting the inclusion criteria. Figure 1 outlines the selection steps.

Fig. 1. Fluxogram of the search strategy



Demographic analysis of studies

In the selected studies, analyzes were carried out in the countries: Italy, Norway, United Kingdom, Sweden, United States, Spain, Denmark, and Hungary. Samples ranged from 1 to 81,337 patients, including 1 case study, 1 case-control study, and 12 cohort studies. Data acquisitions between surveys were predominantly from non-intensive COVID-19 units, hospital records, medical centers, laboratories, and clinics (78.57%); followed by self-reported questionnaires (14.29%); and government data (7.14%).

Symptomatology

It was possible to observe several complaints related to cognitive problems. The neuropsychological complaints cited were: increased forgetfulness (26.7%), increased difficulty concentrating (12% ~ 26.7%), learning difficulties (20% ~ 22%), memory problems (2.8% ~ 82%), fluency verbal (8.4% ~ 34.6%), sleep disorders (30.8% ~ 34.7%) and complaints of neurocognitive impairment (17.8% ~ 58.7%).

Among the articles that passed the inclusion criteria, a case study followed a patient for 8 months after infection and obtained the following neuropsychological complaints: processing speed, logical reasoning, visual memory, subjective neurocognitive complaints, and working memory.

Articles' analysis

Regarding the follow-up of cases, some began while the patients were hospitalized and extended after discharge. All follow-ups after discharge had a mean time ranging from 2 to more than 12 months. General information about the studies can be found in table 1.

Table 1. Included studies.

| Ref | Sample | Follow-up period | Cognitive assessment variables | Main objective | Summary of results |
|------|---------------------------------------|----------------------------------|---|--|---|
| [14] | 45 years old man post severe COVID-19 | 4 and 8 weeks after extubation n | Verbal learning, verbal memory, working memory, visual memory, visuospatial ability, verbal fluency, executive functions, | Elaborate on the possible cognitive sequelae related to COVID-19 | After extubation, M. scored above average in tests of learning, memory, word fluency, and visuospatial functions. Minor deficits were still |

| | | | processing speed, and attention | found in logical reasoning, attention, executive functioning, and processing speed |
|------|--|----------------------------|---------------------------------|---|
| [15] | 445 participants. Of these: 238 symptomatic participants were not hospitalized, and 129 with persistent symptoms for more than 12 weeks. The mean age was 44 for women and 46 for men. | Betwe en days and 3 months | Memory and concentratio n | Determine the prevalence and risk factors for acute and persistent symptoms in non-hospitalized COVID-19 patients. Difficulty with concentration or memory were among the most common persistent (4 to 12 weeks) symptoms (13% of the 198 sample). The risk of persistent symptoms was higher in women and people with high BMI |
| [16] | 312 patients: 247 isolated at home and 65 hospitalized. The mean age was 46. | 2 months | Memory and concentratio n | Assessing the burden of persistent COVID in mild to moderately ill patients 13% of younger participants (0-15 years) rarely experienced persistent symptoms, while 52% of young adults (16-30 years) had persistent symptoms, such as impaired concentration (13%) and memory (11%) |

| | | | | | |
|------|--|------------|---------------------------|--|--|
| [17] | 107 volunteers with a mean age of 63 years (32-90) | 6 months | Conscious ness and memory | Exploring neurological sequelae of SARS-CoV-2 infection | 49 patients complained of the persistence of symptoms (impaired memory, n = 11) |
| [18] | 38 participants aged 22-74 | 4-5 months | Cognitive functions | Study the occurrence of cognitive changes in the months following hospital discharge | 26.7% of participants reported moderate to a severe increase in forgetfulness and lack of concentration; 20% reported difficulty learning new skills or procedures; 26.3% scored low on verbal recall assessment; 10.5% performed low on immediate verbal recall; 18.4% demonstrated impairment in long term visual memory; 15.8% indicated impairment in short term visual memory. Positive correlation between shorter duration of |

| | | | | |
|------|---|--------------------------|---|---|
| | | | | hospitalization and verbal memory consolidation ($r = .40$, $p = .027$) |
| [19] | 100 subjects in the 2nd grade, between 12 and 13 years old | 3 months after infection | Perceptual reasoning, verbal comprehension, working memory, processing speed, higher executive functions, and short-term memory | Understand the consequences of COVID contraction hospitalization about cognitive functioning in preadolescents |
| [20] | Overall sample: 226 survivors (149 men aged 26-87 years, $M = 58.5 +/- 12.8$). Of these, 130 were cognitively assessed | 3 months after discharge | Verbal memory, verbal fluency, working memory, selective attention, processing speed, psychomotor coordination, and executive functions | To study psychopathological and neurocognitive impacts of COVID-19 survivors three months after clinical recovery |
| [21] | 81,337 participants | Unreported | Cognitive functions | Verify the existence of an association between cross-sectional data of the cognitive performance of |

| | | | | | |
|------|---|-----------------------------|--|---|---|
| | | | | the participants | deficits after carefully controlling for premorbid IQ, pre-existing medical conditions, sociodemographic factors, and mental health symptoms |
| [22] | 179 were included in the final analysis | 2 months after recuperation | Immediate verbal memory and learning, Working memory, delayed verbal memory, and verbal fluency | Assess neuropsychiatric and quality of life consequences 2 months after hospital discharge | Considerable prevalence of neurocognitive impairment, low psychiatric comorbidity, and poor QoL in COVID-19 survivors, even in non-critical patients. |
| [23] | 29 participate d in the cognitive evaluation (from a larger initial sample) | 3-4 months | Cognitive functions (verbal memory, working memory, verbal fluency, processing speed, and executive functions) | Investigate the frequency, patterns, and severity of cognitive impairment 3-4 months after hospital discharge, relationship with subjective cognitive complaints, and quality of life | Objective cognitive impairment in 59% to 65% of participants, with greater impairment in verbal learning and executive functions. More than 80% reported severe cognitive difficulties in performing activities of daily living |
| [24] | 101 | 4-12 | Psycholog | To explore | The |

symptomat ic patients, 60 women and 41 men (home care, n = 62, hospitalized, n = 39), Mean age = 50 years months ical fatigue and subjective cognitive impairment possible associations between timing, the severity of fatigue, and memory impairment post-COVID and anti-SARS-CoV-2 antibodies incidence of post-COVID features is significantly higher among women; the strength of the systemic immune response reflected by anti-SARS-CoV-2 antibody titers may have an impact on the severity of post-COVID fatigue and memory impairment.

[25] 165 patients, mean age = 64.8 years 6 months Memory and cognitive functions Investigate whether long-term symptoms and neurological features depend on the severity of COVID-19 High prevalence of post-COVID clinical manifestations, with memory complaints being one of the most prevalent symptoms (31%)

[26] 13,001 participants completed the baseline questionnaire, with a mean age of 47 years. 8 months Memory and confusion Examining self-reported memory problems 8 months after COVID-19 infection In the multiple logistic regression model, SARS-CoV-2 positivity at baseline was strongly associated with reporting memory problems (odds ratio [OR], 4.66;

95% CI, 3.25-6.66). At follow-up, 267 of 649 participants (41%) in the SARS-CoV-2 positive group reported significant worsening of health compared with 1 year earlier, and 81 of 651 participants (12%) in the SARS-CoV-2 - positive group reported concentration problems. In addition, 59 of 267 participants (82%) in the SARS-CoV-2 positive group who reported memory problems also reported worsening health.

| | | | | | |
|------|---|------------|-------------------|---|---|
| [27] | 1,584 recovered COVID-19 patients with mild to severe forms of the disease, aged 18 to 96 years | Unreported | Short-term memory | Test the influence of subjective experiences of discrimination, financial security, and quality of care on disease severity and long-lasting symptom complaints | 25% of participants reported having cognitive or memory problems after COVID-19 infection. The most commonly reported cognitive |
|------|---|------------|-------------------|---|---|

complaints were impairment in short-term memory (70%), attention (58%), and learning (22%).

The level of infection severity and symptoms, according to the studies, are related to age, older patients are more likely to develop a severe state of infection by SARC-CoV-2; with the existence of comorbidities and a high BMI. The studies also showed that in females, post-COVID symptoms lasted longer after hospital discharge.

According to the data described in table 2, the instrument used to assess the studies selected in the research was the Newcastle-Ottawa Scale [13]. Following the scale criteria, 10 cohort studies were rated as good, with 5 of them scoring 66.6%, 3 scoring 77.7%, and 2 scorings 88.9%. Two studies were classified as low quality (55.3% and 33.3%). The case-control study, with a score of 55.6%, was classified as non-standardized. Also included in the research, the case study carried out by [14] could not be evaluated due to a lack of adequate metrics for the type of research.

Table 2. Bias risk assessment.

| Studies | Study design | Quality analysis | | | | Classification |
|---------|-------------------|------------------|---|-----|-----------|-----------------|
| | | S | C | O | Total (%) | |
| [15] | Cohort | +++ | + | ++ | 66.66% | GOOD |
| [16] | Cohort | +++ | + | ++ | 66.66% | GOOD |
| [17] | Cohort | +++ | + | +++ | 77.77% | GOOD |
| [18] | Cohort | +++ | + | ++ | 66.66% | GOOD |
| [20] | Cohort | +++ | + | ++ | 66.66% | GOOD |
| [21] | Cohort | +++ | + | + | 55.6% | POOR QUALITY |
| [22] | Cross - sectional | +++ | + | +++ | 77.77% | |
| [23] | Cohort | +++ | + | ++ | 66.66 % | GOOD |
| [24] | Cohort | +++ | + | +++ | 77.77 % | GOOD |

| [25] | Cohort | ++++ | + | +++ | 88.9% | GOOD |
|---------|--------------|------|---|-----|-----------|---------------------|
| [26] | Cohort | ++++ | + | ++ | 88.9% | GOOD |
| [27] | Cohort | + | - | ++ | 33.3% | POOR QUALITY |
| Studies | Study design | S | C | E | Total (%) | Classification |
| [19] | Case-control | ++ | + | +++ | 55.6% | Not standardized |

Table 2. S = selection; C = comparability; O = outcome; E = exposure.

4 Discussion

This study aimed to bring together the state-of-the-art of knowledge about memory impairment after COVID-19 infection, considering the consequences it may bring to world education in the future. 14 studies were analyzed, including one case report, one case-control, and 12 cohort studies, in which the participants' cognitive impairments were assessed using various neuropsychological instruments.

It is notorious the growth of reports of post-infection neurocognitive deficits, being common the factors of forgetfulness, concentration and learning difficulties, memory impairment, among others. It has been noted that there is a higher incidence of persistent symptoms in female participants [15]. Preliminary causes for the results found are the higher incidence of women in the samples of most studies and the number of women working in the health area, such as nursing technicians, nurses, and physicians, being greater than the number of men [15].

Previous comorbidities, involving cardiorespiratory conditions or individual's BMI, have become important markers to determine the severity of the disease [16, 25]. Patients with high BMI, close to clinical obesity, had symptoms and sequelae that persisted longer at moderate and severe levels [16, 25].

In a systematic review carried out by [34], data from other epidemics and pandemics were retrospectively analyzed to identify the means that a viral infection can bring neurological sequelae. In [23], the use of objective assessments of brain function for the initial identification of neurocognitive symptoms in patients is suggested.

Other studies have already revealed biological aspects, such as the action of the virus on the host's proteases, and the ease of migration from the airways to deeper parts of the brain, causing neurological damage after the treatment of the infection [30, 31, 32]. As well as the influence of hospitalization and medical procedures were shown to be aggravating factors for the development of sequelae [30]. The same result was found in the present study, emphasizing the influence of hospitalization on the outcomes of post-COVID neuropsychological assessments.

In addition to the results of previous studies, a more in-depth analysis of the influence of sex on post-COVID sequelae is needed, as the studies point to a trend of women with lasting sequelae after infection. Moreover, it is necessary to evaluate biological factors that would determine the severity of the sequel, given the non-direct relationship with the age group of the sample.

With the development of cognition sequelae after COVID-19 infection, the need to analyze the consequences for learning emerges. Among the age groups (12 ~ 65 years) affected by these impairments, there are individuals in school, university, and labor market stages.

Applicability solutions for contributing to the transmission of knowledge and memory consolidation include studies in the field of artificial intelligence. Artificial intelligence is a branch of Computer Science that is interested in bringing “intelligent” behaviors and thoughts so that they can be performed by computers [33].

Studies such as [34] bring the idea of using games to improve the teaching-learning process, which would allow the user to improve their skills and provide the construction of knowledge as the individual remains connected to the game [34]. [35] tells us that the use of games can improve the individual's cognitive capacity even though such improvement can rarely be transferred outside the game. A study [36] showed that some games can influence decision-making capacity, as well as changing strategy.

Along with these studies, the use of the term “gamification” addressed by [37] refers to the idea of using games as a basis for mechanics, aesthetics, and thinking in activities. This kind of approach can turn out to be a way to help people with learning difficulties.

The possibility of using interactive educational games brings the idea of personalized education, contributing to the resolution of educational problems [38], in addition to bringing a certain autonomy to the learning process, because through the result of these interactive games, the educator would have access to graphics containing important information to identify how to better direct the learning methods.

Thus, artificial intelligence can be a solution to train memory capacity, improving learning and helping to reduce cognitive impairment after infection by the SARS-CoV-2 virus. With these factors, the application of artificial intelligence provides a wide range of possibilities that can contribute to the continuity of the educational training of young people; adaptation to the reality of the labor market for adult individuals; and cognitive capacity rehabilitation for the elderly.

References

1. Beching, N., Fletcher, T., Robert, F.: COVID-19. BMJ Best Pract (2020).
2. Singhal, T.: A review of coronavirus disease-2019 (COVID-19). Indian J Pediatr. 87, 281-6 (2020).
3. WHO Homepage, <https://covid19.who.int/>, last accessed 2021/10/07.

4. Strunk, J. L., Temesgen, H., Andersen, H., Packalen, P.: Imaging profile of the COVID-19 infection: radiologic findings and literature review authors. 80, 1-8 (2014).
5. Duarte Jr, S. S., Guarnier, G. F. F., Cardoso, I. B. R., Felicio, F. C., Pereira, J. S., Silva, A. C. S. Leitão, R. M.; Neves, M. A. O., Azizi, M. A. A.: Recovery from Post Covid-19 Memory Deficit: a review. Revista de Ciências Biológicas e da Saúde 1, 1-10 (2021).
6. Al-Ramadan, A., Rabab'h, O., Shah, J., Gharaibeh, A.: Acute and Post-Acute Neurological Complications of COVID-19. Neurol. Int. 13, 102–119 (2021).
7. Tang, S. W., Helmeste, D., Leonard, B.: Inflammatory neuropsychiatric disorders and COVID-19 neuroinflammation. Acta Neuropsychiatrica 1, 165-177 (2021).
8. Baddeley, A. D., Anderson, M. C., Eysenck, M. W.: Memória. Artmed, Porto Alegre (2011).
9. Damásio, A. R.: E o Cérebro Criou o Homem. Compainha das Letras, São Paulo (2011).
10. Uversky, V.N., Elrashdy, F., Aljadawi, A., Ali, S.M., Khan, R.H., Redwan, E.M.: Severe acute respiratory syndrome coronavirus 2 infection reaches the human nervous system: How?. J Neurosci Res. 99(3), 750-777 (2021).
11. Moher, D., Liberati, A., Tetzlaff, J., Altman, DG., The PRISMA Group: Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097 (2009).
12. Ouzzani, M., Hammady, H., Fedorowicz, Z., Elmagarmid, A.: Rayyan — a web and mobile app for systematic reviews. Systematic Reviews 5:210 (2016).
13. Wells, G. A., Shea, B., O'Connell, D., Peterson, J., Welch, V., Losos, M., Tugwell, P.: The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomized studies in meta-analyses (2000).
14. Backman, L., Möller, M.C., Thelin, E.P., Dahlgren, D., Deboussard, C., Östlund, G., Lindau, M.: Monthlong Intubated Patient with Life-Threatening COVID-19 and Cerebral Microbleeds Suffers Only Mild Cognitive Sequelae at 8-Month Follow-up: A Case Report. Arch Clin Neuropsychol 17:acab075 (2021)
15. Bliddal, S., Banasik, K., Pedersen, O.B. et al. Acute and persistent symptoms in non-hospitalized PCR-confirmed COVID-19 patients. Sci Rep 11, 13153 (2021).
16. Blomberg, B., Mohn, K.G., Brokstad, K.A., Zhou, F., Linchausen, D.W., Hansen, B.A., Lartey, S., Onyango, T.B., Kuwelker, K., Sævik, M., Bartsch, H., Tøndel, C., Kittang, B.R., Bergen COVID-19 Research Group, Cox, R.J., Langeland, N.: Long COVID in a prospective cohort of home-isolated patients. Nat Med.27(9), 1607-1613 (2021).
17. Bozzetti, S., Ferrari, S., Zanzoni, S., Alberti, D., Braggio, M., Carta, S., Piraino, F., Gabbiani, D., Girelli, D., Nocini, R., Monaco, S., Crisafulli, E., Mariotto, S.: Neurological symptoms and axonal damage in COVID-19 survivors: are there sequelae? Immunol Res. 7, 1–5 (2021).
18. Ferrucci, R., Dini, M., Groppo, E., Rosci, C., Reitano, M.R., Bai, F., Poletti, B., Brugnera, A., Silani, V., D'Arminio, Monforte, A., Priori, A.:

- Long-Lasting Cognitive Abnormalities after COVID-19. *Brain Sci.* 11(2), 235 (2021).
19. Frolli, A., Ricci, M. C., Di Carmine, F., Lombardi, A., Bosco, A., Saviano, E., & Franzese, L.: The Impact of COVID-19 on Cognitive Development and Executive Functioning in Adolescents: A First Exploratory Investigation. *Brain sciences*, 11(9), 1222 (2021).
 20. Mazza, M.G., Palladini, M., De Lorenzo, R., Magnaghi, C., Poletti, S., Furlan, R., Ciceri, F., COVID-19 BioB Outpatient Clinic Study group, Rovere-Querini, P., Benedetti, F.: Persistent psychopathology and neurocognitive impairment in COVID-19 survivors: Effect of inflammatory biomarkers at three-month follow-up. *Brain Behav Immun.* 94, 138-147 (2021).
 21. Hampshire, A., Treder, W., Chamberlain, SR., Jolly, AE., Grant, JE., Patrick, F., Mazibuko, N., Williams, SC., Barnby, JM., Hellyer, P., Mehta, MA.: Cognitive deficits in people who have recovered from COVID-19. *EClinicalMedicine*. 39, 101044 (2021).
 22. Mendez, R., Balanza-Martinez, V., Luperdi, SC., Estrada, I., Latorre, A., Gonzalez-Jimenez, P., Feced, L., Bouzas, L., Yepez, K., Ferrando, A., Hervas, D., Zaldivar, E., Reyes, S., Berk, M., Menendez, R.: Short-term Neuropsychiatric Outcomes and Quality of Life in COVID-19 Survivors. *Journal of Internal Medicine* 290(3), 621-631 (2021).
 23. Miskowiak, K.W., Johnsen, S., Sattler, S.M., Nielsen, S., Kunalan, K., Rungby, J., Lapperre, T., Porsberg, C.M.: Cognitive Impairments Four Months After COVID-19 Hospital Discharge: Pattern, severity, and association with illness variables. *European Neuropsychopharmacology* 46, 39-48 (2021).
 24. Molnar, T., Varnai, R., Schranz, D., Zavori, L., Peterfi, Z., Sipos, D., Tőkés-füzesi, M., Illes, Z., Buki, A., Csecsei, P.: Severe Fatigue and Memory Impairment Are Associated With Lower Serum Level of Anti-SARS-CoV-2 Antibodies in Patients With post-COVID Symptoms. *Journal of Clinical Medicine* 10(19) (2021).
 25. Pilotto, A., Cristillo, V., Piccinelli, S.C., Zoppi, N., Bonzi, G., Sattin, D., Schiavolin, S., Raggi, A., Canale, A., Gipponi, S., Libri, I., Frigerio, M., Bezzi, M., Leonardi, M., Padovani, A.: Long-term Neurological Manifestations of COVID-19: prevalence and predictive factors. *Neurological Sciences* (2021).
 26. Søraas A., Bø R., Kalleberg K.T., Stører N.C., Ellingjord-Dale M., Landrø N.I.: Self-reported Memory Problems 8 Months After COVID-19 Infection. *JAMA network open* 4(7), e2118717 (2021).
 27. Thomason M.E., Hendrix C.L., Werchan D. Brito N.H.: Social Determinants of Health Exacerbate Disparities in COVID-19 Illness Severity and Lasting Symptom Complaints. *medRxiv: the preprint server for health sciences* (2021).
 28. Stefano G. B.: Historical Insight into Infections and Disorders Associated with Neurological and Psychiatric Sequelae Similar to Long COVID. *Medical science monitor: international medical journal of experimental and clinical research*, 27, e931447 (2021).
 29. Valenzano, A., Scarinci, A., Monda, V., Sessa, F., Messina, A., Monda, M., Precenzano, F., Mollica, M. P., Carotenuto, M., Messina, G.,

- & Cibelli, G.: The Social Brain and Emotional Contagion: COVID-19 Effects. *Medicina (Kaunas, Lithuania)*, 56(12), 640 (2020).
30. Łojek, E., Egbert, A., Gambin, M., Gawron, N., Gorgol, J., & Hansen, K. et al.: Neuropsychological disorders after COVID-19. Urgent need for research and clinical practice. *Advances in Psychiatry and Neurology/Postępy Psychiatrii i Neurologii*, 30(2), 104-112 (2021).
31. Mahalakshmi, A. M., Ray, B., Tuladhar, S., Bhat, A., Paneyala, S., Patteswari, D., Sakharkar, M. K., Hamdan, H., Ojcius, D. M., Bolla, S. R., Essa, M. M., Chidambaram, S. B., & Qoronfleh, M. W.: Does COVID-19 contribute to development of neurological disease?. *Immunity, inflammation and disease*, 9(1), 48–58 (2021).
32. Peterson, C. J., Sarangi, A., Bangash, F.: Neurological sequelae of COVID-19: a review. *The Egyptian journal of neurology, psychiatry and neurosurgery*, 57(1), 122 (2021).
33. Lobo, L.C.: Artificial Intelligence, the Future of Medicine and Medical Education. *Revista Brasileira de Educação Médica*, 42, 3 (2018).
34. Kiryakova, G., Angelova, N., and Yordanova, L.: Gamification in education. *Proceedings of 9th International Balkan Education and Science Conference* (2014).
35. Ball, K., Berch, D. B., Helmers, K. F., Jobe, J. B., Leveck, M. D., Marsiske, M., Morris, J. N., Rebok, G. W., Smith, D. M., Tennstedt, S. L., et al.: Effects of cognitive training interventions with older adults: a randomized controlled trial. *Jama* 288(18), 2271–2281 (2002).
36. Agarwal, S., Ram, R., and Iida, H.: Measuring force of game information in the brain: Linking game refinement theory and neuroscience. In *Knowledge and Smart Technology (KST), 2016 8th International Conference on*, 291–294 (2016).
37. Kapp, K.: *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*. Pfeiffer essential resources for training and HR professionals. Wiley (2012).
38. Kay, J.: AI and Education: Grand challenges. *IEEE Intelligent Systems* 27(5), 66– 69 (2012).



III LATIN AMERICAN WORKSHOP ON COMPUTATIONAL NEUROSCIENCE

Proceedings of the III Latin American Workshop on Computational Neuroscience (LAWCN 2021), which addresses: Computational Neuroscience; Neuroscience; Neuroengineering; and Artificial Intelligence and related topics.

Organization



Sponsors and Support

