

## 13º Congresso de Inovação, Ciência e Tecnologia do IFSP - 2022

### APLICAÇÃO DE UMA REDE NEURAL NA PREVISÃO DA EVOLUÇÃO CLÍNICA DE PACIENTES DE COVID-19

THIAGO N. DE SOUSA<sup>1</sup>, MIGUEL ANGELO A. DE SOUSA<sup>2</sup>, RICARDO PIRES<sup>3</sup>, SARA D. DOS SANTOS<sup>4</sup>

<sup>1</sup>Graduando em Engenharia Eletrônica, Bolsista PIBITI, IFSP, Câmpus São Paulo, t.nishimura@aluno.ifsp.edu.br.

<sup>2</sup>Doutor em Ciências, USP, Docente do Departamento de Elétrica, IFSP, Câmpus São Paulo, angelo@ifsp.edu.br.

<sup>3</sup>Doutor em Sistemas Automáticos e Microeletrônicos, Université Montpellier II, Docente do Departamento de Elétrica, IFSP, Câmpus São Paulo, ricardo\_pires@ifsp.edu.br.

<sup>4</sup>Doutora em Ciências, USP, Docente do Departamento de Elétrica, IFSP, Câmpus São Paulo, sarad@ifsp.edu.br.  
Área de conhecimento (Tabela CNPq): 3.13.00.00-6 Engenharia Biomédica.

**RESUMO:** A pandemia de COVID-19 afetou o mundo de forma drástica. Milhões de vidas foram ceifadas e, mesmo com a imunização avançada no Brasil, o vírus ainda se dissemina e sofre mutações. A economia mundial também tem sofrido os impactos das medidas de contenção da crise sanitária, piorando ainda mais o cenário. A comunidade científica está unida em busca de soluções para entender a situação e eliminar (ou minimizar) o problema. Neste cenário, diferentes áreas do conhecimento têm trabalhado de forma colaborativa, como é o caso do presente trabalho. A união das áreas de ciências da saúde e tecnológica visam a obter ferramentas que contribuam no entendimento dessa complexa dinâmica pandêmica. Neste trabalho, uma técnica de aprendizado de máquina foi usada para prever a evolução do quadro clínico de pacientes com COVID-19 a partir dos dados coletados pela instituição parceira. Essa base de dados foi composta por informações de pacientes que deram entrada em hospitais de todo o estado do Espírito Santo e foram utilizadas como elementos dos vetores de entrada da rede neural artificial do tipo Perceptron de Múltiplas Camadas (MLP). A rede construída obteve 82% de acurácia e 78% de sensibilidade.

**PALAVRAS-CHAVE:** aprendizagem de máquina; previsão; COVID-19; inteligência artificial; rede neural artificial.

#### Applying a neural network to predict the outcome of COVID-19 patients

**ABSTRACT:** The COVID-19 pandemic had a drastic impact in the world. Millions of lives were taken and, even with the recent advances regarding vaccination in Brazil, the virus still mutates and spreads. Global economy was also affected due to the policies implemented to contain the sanitary crisis. Academia is united in trying to find solutions to understand and eliminate (or mitigate) the issue. Different fields of knowledge have been working collectively, as in the case of the present paper. The unification of different areas of science and technology have the goal of coming up with tools to contribute to the understanding of this complex dynamic. In this paper, a machine learning technique was used to predict the clinic outcome of COVID-19 patients based on data collected by the partner institution. The data set was composed of information about patients who were hospitalized all across Espírito Santo state, the data being used as feature vectors for the Multilayer Perceptron (MLP) type artificial neural network. The network was able to perform the classification task with 82% accuracy and 78% sensitivity.

**KEYWORDS:** machine learning; prediction; COVID-19; artificial intelligence; artificial neural network.

## INTRODUÇÃO

Dentre as tecnologias passíveis de serem utilizadas na resolução de problemas relacionados à medicina, pode-se destacar a Inteligência Artificial (IA). Em medicina, a IA é utilizada para, através do processamento de grande volume

de dados e algoritmos desenvolvidos por especialistas, propor soluções para questões médicas (LOBO, 2017).

No ano de 2020, o mundo começou a ser atingido pela pandemia da COVID-19. Milhões de pessoas perderam suas vidas e muitas ainda sofrem pelos efeitos que decorrem da enfermidade (FERNANDES et al., 2021).

A COVID-19 é uma doença infecciosa causada pelo vírus SARS-CoV-2. A maior parte dos infectados apresentam sintomas leves. Porém, especialmente para aqueles que possuem comorbidades, os sintomas podem se agravar, podendo haver a necessidade de atenção médica especializada (WHO, 2022).

Neste trabalho será utilizado um tipo de algoritmo chamado de redes neurais artificiais, que recebeu esse nome devido a sua inspiração nos neurônios biológicos do sistema nervoso. A unidade básica de uma rede neural é conhecida como Perceptron. Esse tipo de estrutura, proposta por Rosenblatt, em 1961, é composta por uma camada neural e um neurônio em sua camada de saída (ROSEMBLAT, 1961). Para aplicações mais complexas, como é o caso deste projeto, faz-se necessário um tipo mais elaborado de rede neural artificial chamada Perceptron de Múltiplas Camadas (do inglês, Multilayer Perceptron, MLP).

O objetivo deste trabalho é construir uma rede neural capaz de prever a evolução clínica de pacientes com COVID-19 a partir dos sintomas coletados no momento de entrada no hospital.

## MATERIAIS E MÉTODOS

No desenvolvimento do trabalho foi utilizada a linguagem Python (PYTHON, 2022), com as bibliotecas Pandas (PANDAS, 2022), Matplotlib (MATPLOTLIB, 2022), Numpy (NUMPY, 2022) e Tensor Flow(GOOGLE, 2022b). Os códigos foram desenvolvidos na plataforma Google Colab. (GOOGLE, 2022a)

Primeiramente, foi feita a análise exploratória dos dados, visando a selecionar quais colunas seriam utilizadas para compor os vetores de características da rede neural. Essa análise foi realizada no programa Excel e a planilha compartilhada não continha informações que pudessem identificar os pacientes ou algum dado sensível. Foi decidido, em conjunto com o colaborador do trabalho, prof. Dr. Valerio Barauna, da Universidade Federal do Espírito Santo (UFES), que seriam utilizados os sintomas para compor as entradas (23 no total) da rede e a evolução clínica (óbito ou cura) como saída. Os sintomas, que variavam desde febre, perda do olfato, dor no corpo até situações menos convencionais, tais como, conjuntivite ou coma, foram codificados seguindo a lógica One-Hot, que consiste em transformar cada dado em uma variável binária com apenas um bit igual a 1 (MÜLLER; GUIDO, 2016).

Definidos os parâmetros de entrada e saída, foi realizado o treinamento da rede neural. A topologia foi escolhida comparando e treinando redes com diversas quantidades de neurônios nas diversas camadas do MLP, começando com a rede mais básica, ou seja, sem camadas intermediárias (23-1) e, depois, testando diferentes arquiteturas com quantidades diferentes de camadas intermediárias.

Para avaliar o desempenho da rede neural, foram utilizados parâmetros como acurácia, sensibilidade e especificidade. Esses parâmetros foram propostos pelo matemático Thomas Bayes e têm por função avaliar o desempenho das previsões feitas pela rede neural (KAWAMURA, 2002). A figura 1 mostra um infográfico do processo de desenvolvimento do trabalho desde a aquisição até a análise dos dados, a figura foi feita com a auxílio da plataforma Diagrams.net (JGRAPH, 2022).



Figura 1: Infográfico do desenvolvimento do trabalho.

As métricas de interesse avaliadas neste projeto foram a acurácia (1), a sensibilidade (2) e a especificidade (3).

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$Sensibilidade = \frac{VP}{VP + FN} \quad (2)$$

$$Especificidade = \frac{VN}{VN + FP} \quad (3)$$

em que, VP - Verdadeiros Positivos; VN - Verdadeiros Negativos; FP - Falsos Positivos; FN - Falsos Negativos.

A acurácia é o coeficiente que expressa a quantidade de acertos em relação ao total de testes. Essa métrica serve como parâmetro inicial de desempenho do algoritmo.

Neste projeto, o ideal é aumentar a sensibilidade, ou seja, diminuir a quantidade de falsos negativos pois, pensando em uma aplicação hospitalar, os casos de falsos positivos poderiam implicar em uma atenção médica mais acentuada, porém nenhum risco iminente ao paciente. Já os falsos negativos poderiam ocasionar a alta equivocada de um paciente e o possível agravamento do seu quadro clínico longe do hospital, o que seria indesejável. Por essa razão, procurou-se otimizar a sensibilidade da rede ao longo dos diferentes testes realizados.

## RESULTADOS E DISCUSSÃO

Os primeiros resultados obtidos foram os que definiram a topologia da rede neural. Para isso comparou-se o treinamento de diferentes topologias com 23 entradas e uma saída, começando com a rede mais básica, ou seja, sem camadas intermediárias. Após isso foram testadas diferentes arquiteturas com quantidades diferentes de camadas intermediárias, conforme a tabela 1.

Topologia	23-1	23-2-1	23-6-1	23-10-1	23-14-1	23-18-1	23-22-1
Acurácia	81%	81%	82%	82%	82%	82%	82%
Sensibilidade	76%	74%	76%	76%	78%	76%	76%

Tabela 1: Tabela de desempenho das redes testadas.

Dessa maneira, com o intuito de maximizar a sensibilidade e visando deixar a rede com a menor quantidade de neurônios (pensando principalmente na velocidade de processamento), obteve-se que a topologia 23-14-1 (ou seja, 23 neurônios na camada de entrada, 14 neurônios na camada intermediária “escondida” e 1 neurônio na saída) apresentou desempenho levemente superior à anterior, com 78% de sensibilidade e 82% de acurácia.

A figura 2 representa a arquitetura da rede neural escolhida, que foi desenvolvida com o auxílio do programa Multiple Back Propagation (LOPES, 2022).

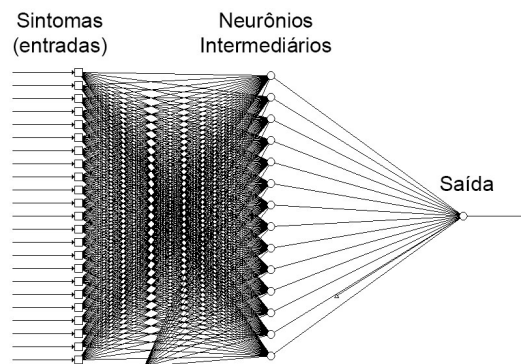


Figura 2: Representação da arquitetura da rede neural 23-14-1.

A figura 3 mostra a matriz de confusão da rede neural construída. Essa matriz traz os resultados dos testes realizados pela rede, ou seja, VP,VN,FP,FN.

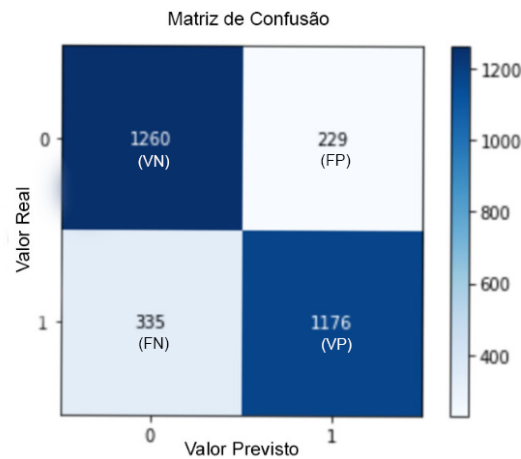


Figura 3: Matriz de confusão gerada à partir da rede neural 23-14-1.

## CONCLUSÕES

A rede neural construída foi capaz de realizar a previsão da evolução clínica dos pacientes com satisfatória acurácia e sensibilidade, atingindo, respectivamente, 82% e 78% em cada caso. Diferentes arquiteturas foram testadas de forma que a rede MLP com 23 neurônios na camada de entrada, 14 neurônios na camada intermediária e 1 neurônio na camada de saída foi aquela que apresentou o melhor desempenho. Como dados de entrada da rede neural, foram usados 23 sintomas dos pacientes com COVID 19 e, como saída da rede, a evolução clínica de cada caso para cura ou óbito. Dessa forma, uma possível continuidade do presente projeto é a identificação dos sintomas mais relevantes para o processo de aprendizagem da rede neural, de forma a identificar os marcadores mais relevantes para o objetivo proposto.

## AGRADECIMENTOS

Ao CNPq pela bolsa de iniciação científica e ao professor Dr. Valerio Baraúna, da UFES, pelo compartilhamento dos dados.

## REFERÊNCIAS

- FERNANDES, P. et al. Life post-covid-19: symptoms and chronic complications. *São Paulo Medical Journal*, São Paulo, Brasil, v. 139, n. 1, 2021. Disponível em: <<https://www.scielo.br/j/spmj/a/vqXLGfYXRvYWcQPFqkYtQsH/?lang=en>>.
- GOOGLE. *Google Colab*. EUA, 2022. Disponível em: <<https://colab.research.google.com/>>. Acesso em: 15 Setembro. 2022.
- GOOGLE. *Tensor Flow*. EUA, 2022. Disponível em: <<https://www.tensorflow.org/overview?hl=pt-br>>. Acesso em: 15 Setembro. 2022.
- JGRAPH. *Diagrams.net*. EUA, 2022. Disponível em: <<https://www.diagrams.net/>>. Acesso em: 17 Setembro. 2022.
- KAWAMURA, T. Interpretação de um teste sob a visão epidemiológica. *Arquivos Brasileiros de Cardiologia*, Brasil, v. 79, n. 4, p. 420-421, 2002. Disponível em: <[#](https://www.scielo.br/j/abc/a/H5MFfM6Syr6HwLnNPdwRqyg/?lang=pt)>.
- LOBO, L. C. Inteligência artificial e medicina. *Revista Brasileirade Educação Médica*, Brasil, v. 41, n. 2, 2017. Disponível em: <<https://www.scielo.br/j/rbem/a/f3kqKJjVQJxB4985fDMVb8b/?lang=pt>>.
- LOPES, N. *Multiple Back Propagation*. [S.l.], 2022. Disponível em: <<http://mbp.sourceforge.net/contact.html>>. Acesso em: 17 Setembro. 2022.

MATPLOTLIB. *Matplotlib 3.6.0 documentation*. [S.l.], 2022. Disponível em: <<https://matplotlib.org/stable/index.html>>. Acesso em: 17 Setembro. 2022.

MÜLLER, A.; GUIDO, S. *Introduction to Machine Learning with Python: A Guide for Data Scientists*. O'Reilly Media, 2016. ISBN 9781449369897. Disponível em: <<https://books.google.com.br/books?id=vbQIDQAAQBAJ>>.

NUMPY. *NumPy documentation*. [S.l.], 2022. Disponível em: <<https://numpy.org/doc/stable/>>. Acesso em: 17 Setembro. 2022.

PANDAS. *Pandas documentation*. [S.l.], 2022. Disponível em: <<https://pandas.pydata.org/docs/>>. Acesso em: 17 Setembro. 2022.

PYTHON. *Python 3.10.7 documentation*. [S.l.], 2022. Disponível em: <<https://docs.python.org/3/>>. Acesso em: 17 Setembro. 2022.

ROSEMBLAT, F. Principles of neurodynamics perceptrons and the theory of brain mechanisms. *Cornell Aeronautical Lab Inc*, Buffalo, NY, EUA, 1961.

WHO. *Coronavirus disease (COVID-19)*. [S.l.], 2022. Disponível em: <[https://www.who.int/health-topics/coronavirus/#tab=tab\\\_1](https://www.who.int/health-topics/coronavirus/#tab=tab\_1)>. Acesso em: 07 Setembro. 2022.