

Análise de pré-condições existentes em pacientes com COVID-19: aplicação de redes neurais na segmentação de pacientes

Alexandre Ferreira Oliveira*, Letícia Martins Raposo **, Diogo Antonio Tschoeke***,
Flávio Fonseca Nobre****, Antonio Mauricio M. de Sá *****.

*Programa de Engenharia Elétrica - COPPE – UFRJ

BRASIL (Tel: (21) 3938-8205; e-mail: alexandre.oliveira@coppe.ufrj.br).

** Programa de Engenharia Biomédica - COPPE/UFRJ (e-mail: raposo@peb.ufrj.br).

*** Programa de Engenharia Biomédica - COPPE/UFRJ (e-mail: diogoat@peb.ufrj.br).

**** Programa de Engenharia Biomédica - COPPE/UFRJ (e-mail: flavio@peb.ufrj.br).

***** Programa de Engenharia Biomédica - COPPE/UFRJ (e-mail: amflms@peb.ufrj.br).

Abstract: In recent years, Artificial Intelligence (AI) has advanced in diverse science fields. SARS Cov-2 is a virus of coronaviridae family that can affect different parts of the human body, being quickly spread, which pushed WHO to decree the pandemic, with severe impacts on global scale. This work aims to employ artificial neural networks in the process of classifying the risk of death of patients, based on pre-existing conditions. We have developed Multilayer Perceptron (MLP) models that, based on patient characteristics, such as age, sex, time of infection, respiratory diseases, etc. predict with accuracy greater than 85% which patients are at risk of dying. Additionally, we compared the proposed model with the Support Vector Machine (SVM) method, which presented results slightly inferior to neural networks.

Resumo: O emprego da Inteligência Artificial (IA) tem ganhado importância nas mais diversas áreas da ciência. O Sars-Cov-2 é um vírus da família coronaviridae que pode afetar diferentes partes do corpo humano, tendo se disseminado rapidamente, o que fez com que a OMS decretasse o estágio de pandemia, com severos impactos em escala global. Este trabalho tem por objetivo construir e empregar redes neurais artificiais no processo de classificação de risco de morte de pacientes, baseado em condições pré-existentes. Modelos perceptron multicamadas (MLP, do inglês, *MultiLayer Perceptron*) foram desenvolvidos a partir de características de pacientes com COVID-19, tais como idade, sexo, tempo de infecção, doenças respiratórias, etc. Os modelos apresentaram acurácia acima de 85% na predição de morte por pacientes com COVID-19. Adicionalmente, os modelos propostos foram comparados ao método de *Support Vector Machine* (SVM), que apresentou resultados ligeiramente inferiores às redes neurais.

Keywords: COVID-19; Diagnosis; MLP; Pre-conditions; Neural Network; SVM.

Palavras-chaves: COVID-19; Diagnóstico; MLP; Pré-condições; Rede Neural; SVM.

1. INTRODUÇÃO

O emprego de redes neurais na predição de resultados em medicina e suporte a diagnósticos encontra grande amparo científico e vem se ampliando continuamente (Trujillano, et al., 2004) (Al-Shayea, 2011).

A rápida disseminação do novo coronavírus Sars-CoV-2 (Gorbalenya, et al., 2020) deu origem a uma doença designada por COVID-19 pela Organização Mundial de Saúde (OMS), que apresenta um espectro clínico variando de infecções assintomáticas a quadros graves. Aproximadamente 20% dos pacientes requerem atendimento hospitalar por apresentarem dificuldade respiratória, e cerca de 5% deles podem necessitar de suporte ventilatório (Saúde, 2020).

O real impacto da COVID-19 é incerto, entretanto, a doença colapsou os sistemas de saúde em distintos locais, comprometendo a oferta de recursos humanos e outros

insumos essenciais para o funcionamento adequado dos serviços de saúde. Nesse contexto, todos os recursos que possam ser empregados para melhorar os processos de suporte a diagnóstico são importantes para ajudar a diminuir os efeitos sobre todo o sistema de saúde.

Neste estudo, usando a metodologia de aprendizado de máquina (ML, do inglês, *machine learning*), aplicamos os modelos *MultiLayer Perceptron* (MLP) com 2 camadas escondidas, MLP com 1 camada escondida e *Support Vector Machine* (SVM) a fim de identificar, dentre os pacientes que fazem uso do sistema de saúde, aqueles que possuem um maior risco de morte.

Para a elaboração do trabalho foram utilizados dados públicos disponibilizados pelo Governo Mexicano, compilados e compartilhados por meio da plataforma Kaggle (Mukherjee, 2020). O banco de dados continha dados anonimizados dos pacientes, informações de sexo, idade e condições de doenças pré-existentes. Assim, o presente trabalho pretende apoiar, por

meio de ferramentas de ML, o processo de triagem de pacientes com COVID-19 em situação de risco.

2. TRABALHOS RELACIONADOS

Em trabalho recente, Brito e Oliveira (2020), a partir de uma base de dados disponível no Kaggle (acessível por meio do link: <https://bit.ly/2YjEqfa>), utilizaram um classificador MLP na análise de características para o auxílio no diagnóstico da COVID-19. Eles aplicaram o método *Sequential Feature Selection* (SFS) que inclui, gradativamente no modelo, variáveis preditoras a partir de um conjunto de variáveis candidatas, até que a adição de uma nova variável não leve a uma redução significativa da função objetivo. Os autores utilizaram uma base de dados de 86 atributos e 395 observações divididas em duas classes para o teste de COVID-19. Eles identificaram seis características como significativas: paciente internado em enfermaria, infectado com influenza A (teste rápido), quartil da idade do paciente, taxas normalizadas de hemoglobina, leucócitos e saturação de oxigênio.

Um estudo de Ahamad et al. (2020), baseado em dados coletados de pacientes da região de Wuhan/Hubei (China) e contendo informações sobre sexo, idade, residência habitual, trabalho, histórico de contato, viagens e histórico de sintomas/doenças (febre, tosse, dor muscular, pneumonia, diarreia, coriza), identificou que o histórico de infecção pulmonar, tosse, febre, idade e histórico de viagens foram as características mais associadas ao status positivo de COVID-19.

Mei et al. (2020) aplicaram no diagnóstico da COVID-19 algoritmos de inteligência artificial (IA) para integrar os dados gerados pela tomografia computadorizada (TC) de tórax aos sintomas clínicos, histórico de exposição e testes laboratoriais. Quando as tomografias e a história clínica associada estão disponíveis, a rede MLP proposta pode ajudar a diagnosticar rapidamente os pacientes com COVID-19.

Um estudo de Rocha et al. (2020) empregou redes neurais artificiais MLP e *Radial Basis Function* (RBF) na previsão de contágio e óbitos por COVID-19 para o estado do Pará, que acabou indicando a necessidade de serem adicionados 1.000 leitos de UTI à rede hospitalar do estado.

3. MÉTODO DE PESQUISA

3.1 Base de Dados

Para o treinamento dos modelos utilizados nesse trabalho, foram considerados os dados públicos disponibilizados pelo Governo Mexicano, compilados e compartilhados através da plataforma Kaggle (Mukherjee, 2020). A base de dados possui informações de 566.602 pacientes atendidos pelo sistema de saúde daquele país. A partir da base original, foram derivadas três variáveis utilizadas nas análises. A Tabela 1 apresenta todas as variáveis utilizadas nesse estudo.

Variável	Tipo	Descrição
id	Char	Id do paciente
sex	Integer	Sexo
age	Integer	Idade
patient_type	Integer	Atendido em enfermaria ou internado
icu	Integer	Se deu entrada em CTI
intubed	Integer	Se foi entubado
entry_date	Date	Data em que entrou para o atendimento
date_symptoms	Date	Data em que manifestou sintomas
date_died	Date	Data em que morreu
pneumonia	Integer	Se teve pneumonia
pregnancy	Integer	Se está grávida
diabetes	Integer	Se tem diabetes
copd	Integer	Se tem obstrução pulmonar crônica
asthma	Integer	Se tem asma
inmsupr	Integer	Se é imunodeprimido
hypertension	Integer	Se é hipertenso
other_disease	Integer	Se possui outras doenças
cardiovascular	Integer	Se tem problemas cardiovasculares
obesity	Integer	Se é obeso
renal_chronic	Integer	Se tem doença renal crônica
tobacco	Integer	Se é tabagista
contact_other_covid	Integer	Se teve contato com pessoa com COVID-19
covid_res	Integer	Se fez teste para COVID-19
Dt_Sintoma_Entrada (Derivada)	Integer	Diferença entre a data de sintomas e entrada para tratamento
Dt_Sintoma_Morte (Derivada)	Integer	Diferença entre as datas de sintomas e morte
Morreu (Derivada)	Binary	Indica se paciente morreu

Tabela 1 - Variáveis do estudo.

3.2 Equipamento e Ferramentas

Neste trabalho, o hardware utilizado foi um notebook i7-6500U @ 2.5GHz-2.59GHz, com 16GB, rodando Windows 10.

Foi utilizada a plataforma Anaconda 1.9.12, com Jupyter Notebook, usando Python 3.6.12, biblioteca Keras 1.1.2, TensorFlow-CPU 2.3.1 e TensorBoard 2.3.0.

3.3 Pré-Processamento

Antes de iniciar o treinamento das redes, foi realizado um pré-processamento do banco de dados, seguindo o proposto por Haykin (2009). Foi gerada uma base de dados com quantidade de ocorrências balanceadas para a variável de saída. A base de

dados inicial continha informações de 566.602 pacientes, dos quais 36.172 haviam morrido e 530.426 não haviam morrido. Portanto, fez-se necessário o balanceamento de base antes do processo de treinamento das redes, com a escolha aleatória de 36.172 de pacientes que não haviam morrido, caso contrário a rede tenderia a classificar todos os pacientes com baixo risco de morte. A Tabela 2 apresenta as dimensões e propósito da base:

Propósito BD	Variáveis	Saída	Casos
Identificar pacientes com maior risco de morte (A)	20	Morreu (S/N)	72.344

Tabela 2 - Propósito da Base de Dados.

3.4 Treinamento e Teste

Para o processo de treinamento, a base foi dividida em 3 subconjuntos, a saber: Treinamento, Validação e Teste, respeitando a proporção de 60%, 20% e 20%, respectivamente.

3.5 Indicadores

Durante o processo de treinamento das MLPs foram obtidos os dados da acurácia de treinamento, acurácia de validação, erro de treinamento e erro de validação. Cada rodada de treinamento possuía 200 épocas, com a gravação dos parâmetros do melhor modelo e parada prematura (*early stopping*) em um intervalo de 50 épocas.

Para cada um dos modelos simulados, foi elaborada uma matriz de confusão (Figura 1), para que fossem extraídos os indicadores apontados a seguir.

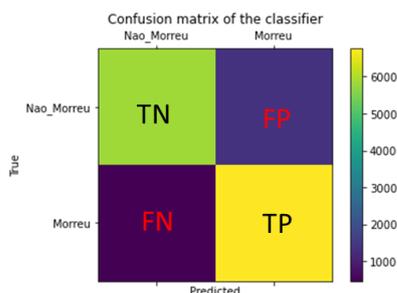


Figura 1 - Exemplo de Matriz de Confusão sendo TN: Verdadeiro Negativo (True Negative), TP: Verdadeiro Positivo (True Positive), FN: Falso Negativo (False Negative) e FP: Falso Positivo (False Positive).

- Acurácia (ACC) - Mede a razão entre os casos classificados corretamente e o número total de casos (Agarwal, 2019).

$$ACC = \frac{TN + TP}{TP + TN + FP + FN}$$

- Precisão (PRE) - Mede a razão entre os casos verdadeiramente positivos e o total de positivos apontados

pelo modelo (Garate-Escamila, et al., 2020) sendo também conhecido por Valor Preditivo Positivo.

$$PRE = \frac{TP}{TP + FP}$$

- Recall (REC) - Mede a razão entre os casos verdadeiramente positivos e a soma dos verdadeiros positivos com os falsos negativos (Agarwal, 2019), sendo também conhecido por Sensibilidade.

$$REC = \frac{TP}{TP + FN}$$

- F1 - Representa a média harmônica da precisão e do recall (Agarwal, 2019) (Garate-Escamila, et al., 2020).

$$F1 = 2 \left(\frac{Precision \times Recall}{Precision + Recall} \right)$$

- Especificidade (ESP) - Mede a razão entre os casos classificados corretamente como negativos e a soma dos verdadeiramente negativos com os falsos positivos (Agarwal, 2019).

$$ESP = \frac{TN}{TN + FP}$$

3.5 Processo de Pesquisa

Inicialmente, foi efetuada uma análise exploratória de dados por meio da análise dos gráficos de barras das variáveis categóricas, onde não se observou qualquer anomalia. Posteriormente, foram elaborados histogramas e gráficos *box plot* para as variáveis idade e tempo de sintomas, tendo sido eliminadas três observações que apresentavam data da morte anterior ao início dos sintomas. Também foi descartada a variável DT_Sintoma_Morte, pois a mesma estava 100% relacionada à variável de saída (possuía valor 0 para pacientes que não morreram e um valor inteiro para aqueles que morreram).

Usando como referência a base de dados A (Figura 2), foi construída uma rede MLP com a primeira camada contendo três neurônios e a segunda camada com dois neurônios, com função de ativação tangente hiperbólica. Esta função é não linear contínua com intervalo na faixa de [-1,1] e adequada para camadas escondidas (Haykin, 2009). Por outro lado, na camada de saída foi utilizada a função de ativação sigmoide cujos valores de saída estão na faixa de 0 a 1, sendo adequada para fornecer valores facilmente associáveis à probabilidade de determinada instância pertencer ou não a determinada classe (Agarwal, 2019).

No primeiro conjunto de modelos testados, diferentes otimizadores foram utilizados para verificar a diferença de tempo no processo de treinamento. Nessa etapa, foram usados os métodos de otimização de gradiente ADAM, SGD com taxas de aprendizado iguais a 0,01 e 0,5, respectivamente. Os resultados foram de aproximadamente sete, doze e treze minutos, respectivamente (Figura 2 - T1a, T1b, T1c).

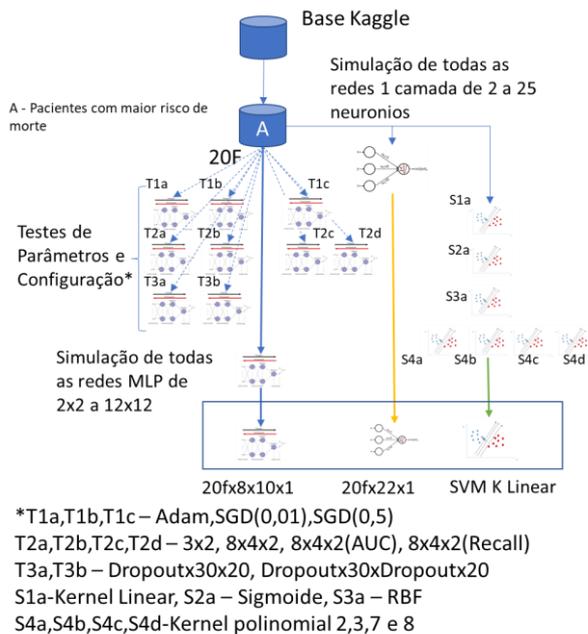


Figura 2- Processo de pesquisa, onde de T1a a T1c foram avaliados os métodos de otimização de gradiente, de T2a a T2d diferentes funções de ativação, de T3a a T3b diferentes configurações de dropout. Diferentes funções de kernel foram avaliadas de S1a a S4d. Simulações da MLP de 2 camadas, variando de 2 a 12 neurônios, por camada, e a MLP de 1 camada variando de 2 a 25 neurônios. A topologia MLP de 2 camadas com 8x10 neurônios, a MLP de 1 camada com 22 neurônios e a SVM com kernel linear foram as que apresentaram os melhores resultados.

No segundo conjunto de modelos testados, os tipos de função de ativação foram mantidos, mas o número de camadas foi alterado e as métricas, avaliadas. Os valores do conjunto de métricas testadas não sofreram grandes diferenças entre as distintas configurações (Figura 2 – T2a, T2b, T2c, T2d).

No terceiro conjunto de modelos avaliados, a rede foi alterada para uma primeira camada com 30 neurônios e uma segunda camada com 20 neurônios e um neurônio na camada de saída. Para essa rede, foi inicialmente aplicado um *dropout* de 0,5, que consiste em omitir aleatoriamente neurônios durante o processo de treinamento para prevenir o sobreajuste (*overfitting*) (Srivastava, et al., 2014) antes da primeira camada. Essa abordagem reduziu a acurácia para 77,6%. Uma segunda simulação incluindo mais um *dropout* de 0,5 antes da segunda camada reduziu a acurácia para 40,6% (Figura 2 – T3a, T3b).

Após essas etapas, variáveis preditoras foram avaliadas a fim de verificar se alguma delas deveria ser descartada antes do processo de simulação final. As 20 variáveis foram mantidas.

A partir desse ponto, simulações usando como parâmetros de entrada os melhores resultados obtidos na etapa de teste e configuração - rede MLP com duas camadas variando de dois a doze neurônios por camada e camada de saída com um neurônio, otimizador ADAM, sem *dropout* e métrica acurácia - foram realizadas (Figura 2 - T1a ... T3b).

A rede MLP identificada com melhor acurácia na base de validação apresentou 20 variáveis de entrada, primeira camada com oito neurônios, segunda camada com dez neurônios e camada de saída com um neurônio.

Posteriormente, foi efetuado um novo conjunto de simulações para uma MLP de apenas uma camada, variando o número de neurônios de duas a 25 unidades. A MLP que alcançou o melhor resultado possuía 22 neurônios.

Foi feito também um conjunto de simulações usando o modelo SVM, variando as funções kernel (linear, sigmoide, RBF (*radial basis function*) e polinomial de grau 2, 3, 7 e 8) (Figura 2 – S1a, S2a, S3a, S4a-S4d). O melhor resultado foi obtido com o *kernel* linear

Por fim, foram feitas 30 simulações, de cada um dos dois melhores modelos, quais sejam, MLP de duas camadas (20f x 8x10x1) e MLP de uma camada (20f x 20x1) a fim de se comparar com o resultado do modelo SVM de *kernel* linear.

4. RESULTADOS

O estudo teve por objetivo classificar o risco de morte para pacientes com COVID-19 a partir de variáveis clínicas e de comorbidades.

A rede MLP desenvolvida com 20 variáveis na camada de entrada, oito neurônios na primeira camada escondida, dez neurônios na segunda camada escondida e um neurônio na camada de saída (20x8x10x1) foi avaliada a partir de um conjunto de teste com 14.469 casos. Os resultados estão apresentados na Figura 3. Percebe-se que os valores da diagonal principal da matriz de confusão são de 5 a 10 vezes maiores que aqueles localizados fora da diagonal principal. A acurácia obtida foi de 87,4% e dentre os indicadores apresentados, destaca-se o resultado do *recall*, igual a de 93,10%, representando uma capacidade bem elevada do modelo de não gerar falsas previsões negativas. Como pode-se observar, os resultados foram bastantes satisfatórios, suscitando a possibilidade de replicar esse tipo de rede no contexto do Brasil.

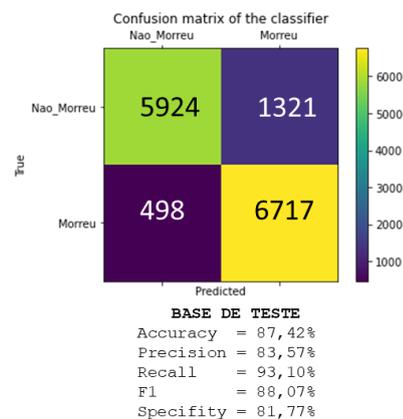


Figura 3 - Resultados da MLP em que se destaca o recall com 93,10%, ressaltando a adequação do modelo em maximizar a razão entre os casos verdadeiramente positivos e a soma de verdadeiros positivos e falsos negativos.

A Figura 4 apresenta os resultados obtidos a partir do conjunto de teste para a MLP contendo 20 variáveis na camada de entrada, 22 neurônios na primeira camada e um neurônio na camada de saída (20x22x1). Como pode ser observado, os resultados de todos os indicadores avaliados foram praticamente iguais aos apresentados pela rede MLP de duas camadas. O *recall* para essa rede alcançou o valor de 93,53%.

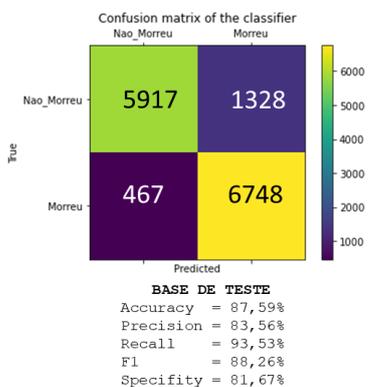


Figura 4 - Resultados da MLP de uma camada onde se destaca o recall com 93,53% com resultado ligeiramente superior a MLP de duas camadas.

A Figura 5 apresenta os resultados para o modelo SVM com *kernel* linear. Como pode-se observar, os resultados dos indicadores foram superiores a 80%, tendo, por exemplo, uma especificidade de 80,19%. Entretanto, esse modelo foi o que apresentou, comparativamente resultados inferiores aos dos demais modelos.

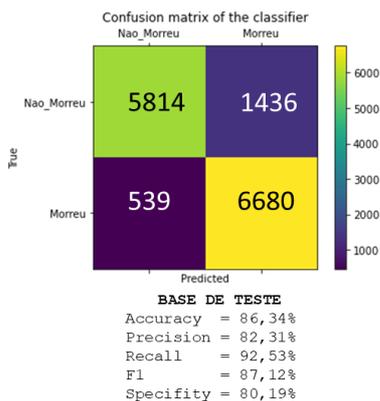


Figura 5 - Resultados da SVM que apesar de ter apresentado um recall de 92,53% foi o modelo com o pior desempenho dos três analisados.

Por fim, com o objetivo de avaliar os impactos dos parâmetros de inicialização na variabilidade dos resultados obtidos para MLP de duas camadas (20x8x10x1) e MLP de uma camada escondida (20x22x1), foram efetuadas 30 simulações

completas para cada uma das topologias, para os quais apresentamos os resultados na Tabela 3, Tabela 4 e no *boxplot* da Figura 6. Como pode-se observar, as redes neurais apresentaram resultados equivalentes entre si, considerando a margem de erro, contudo melhores que o modelo SVM.

MLP 2C	Média	Mediana	Desvio Padrão
Acurácia	87,29%	87,30%	0,07%
Precisão	83,35%	83,43%	0,28%
Recall	93,13%	93,10%	0,44%
F1	87,97%	87,98%	0,08%
Especificidade	81,48%	81,57%	0,46%

Tabela 3 - Resultados dos indicadores para a MLP com duas camadas escondidas.

MLP 2C	Média	Mediana	Desvio Padrão
Acurácia	87,43%	87,41%	0,06%
Precisão	83,67%	83,69%	0,13%
Recall	92,94%	92,94%	0,26%
F1	88,06%	88,05%	0,07%
Especificidade	81,94%	81,97%	0,22%

Tabela 4 - Resultados dos indicadores para a MLP com uma camada escondida.

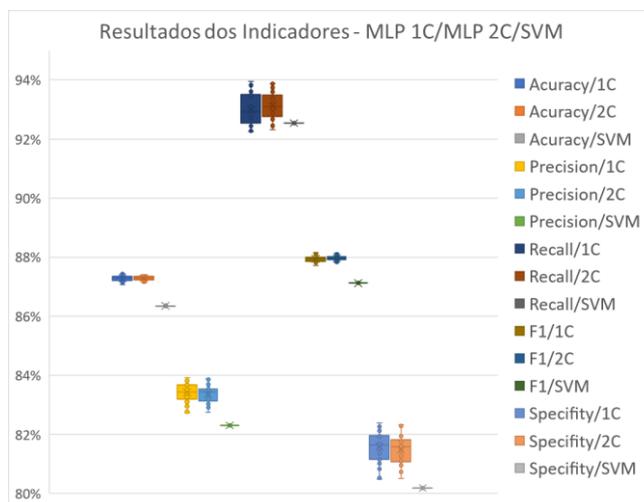


Figura 6 - Gráfico com os resultados das simulações da MLP com uma camada escondida, MLP com duas camadas escondidas e SVM, onde observa-se que as redes neurais apresentaram resultados equivalentes entre si, considerando a margem de erro, contudo melhores que o modelo SVM.

5. CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho efetuou um estudo do emprego de redes MLP e SVM em dados reais de pacientes com COVID-19, o que o torna relevante no atual cenário de pandemia.

Os modelos desenvolvidos foram capazes de prever com uma acurácia maior que 86% quais pacientes possuem risco de morte, tendo as redes MLP de uma ou duas camadas apresentado desempenhos superiores ao modelo SVM, portanto, mostrando resultados bastante animadores e que poderiam ser empregados no cenário brasileiro.

Como trabalhos futuros, sugerimos a aplicação dessa metodologia para diferentes recortes da base, de forma a permitir a identificação dos pacientes com o maior risco de serem intubados e de identificar, dentro do grupo de pacientes que utilizaram UTI, aqueles com maior risco de morte. Adicionalmente, a fim de melhorar o desempenho do modelo de proposto, sugerimos que sejam testadas configurações de redes MLP com três camadas, além de ser efetuado um estudo longitudinal de acompanhamento dos pacientes, de forma a mensurar o impacto dos fatores de risco em diferentes momentos de vida dos mesmos.

AGRADECIMENTOS

Os autores gostariam de agradecer ao CNPq e à FAPERJ pelo apoio a este trabalho. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES) – Código de Financiamento 001.

REFERÊNCIAS

- Agarwal, R., 2019. *The 5 Classification Evaluation metrics every Data Scientist must know*. [Online] Available at: <https://towardsdatascience.com/the-5-classification-evaluation-metrics-you-must-know-aa97784ff226>. [Acesso em 19/12/2020].
- Ahamad, M. et al., 2020. A machine learning model to identify early stage symptoms of SARS-Cov-2 infected patients. *Expert Systems with Applications*, 160(113661).
- Al-Shayea, Q. K., 2011. Artificial Neural Networks in Medical Diagnosis. *International Journal of Computer Science*, 8(2), pp. 150-154.
- Bishop, C. M., 2006. *Pattern Recognition and Machine Learning*. Cambridge: Springer.
- Boser, B. E., Guyon, I. M. & Vapnik, V. N. A., 1992. *Training Algorithm for Optimal Margin Classifiers*. Pittsburgh, s.n.
- Brito, R. & Oliveira, A., 2020. *Análise de Características a Partir do Classificador MLP Para Auxílio no Diagnóstico da COVID-19*. [Online] Available at: <https://doi.org/10.5753/ercemapi.2020.11472>. [Acesso em 18/12/2020].
- Burges, C., 1998. A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery*, 2(doi.org/10.1023/A:1009715923555), pp. 121-167.
- Ceccon, D., 2020. *Funções de ativação: definição, características, e quando usar cada uma*. [Online] Available at: <https://iaexpert.academy/2020/05/25/funcoes-de-ativacao-definicao-caracteristicas-e-quando-usar-cada-uma/> [Acesso em 06/01/2021].
- Evsukoff, A., 2020. *Inteligência computacional : Fundamentos e aplicações*. Rio de Janeiro: e-papers.
- Garate-Escamila, A. K., Hassani, H. & Andres, E., 2020. Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, 19(100330).
- Gorbalenya, A. E. et al., 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology*, Volume 5, pp. 536-544.
- Haykin, S., 2009. *Neural Networks and Learning Machines*. Hamilton: Pearson Prentice Hall.
- Kingma, D. P. & Ba, J. L., 2015. *Adam: a Method for Stochastic Optimization*. San Diego, s.n.
- Mei, X., Lee, H.-C., Yan, Y. & et., a., 2020. Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nature Medicine*, Volume 26, p. 1224-1228.
- Mukherjee, T., 2020. *COVID-19 patient pre-condition dataset*. [Online] Available at: <https://www.kaggle.com/tanmoynx/covid19-patient-precondition-dataset> [Acesso em 18/12/2020].
- Ribeiro, A. M. & Junior, F. d. P. S. A., 2020. Um Estudo Comparativo Entre Cinco Métodos de Otimização Aplicados Em Uma RNC Voltada ao Diagnóstico do Glaucoma. *Revista de Sistemas e Computação*, 10(1), pp. 122-130.
- Rocha, J. et al., 2020. Redes Neurais Artificiais na Previsão de Contágio e Óbitos por COVID-19: Um Estudo no Estado de Pará, BRASIL. *International Journal of Development Research*, 10(4), pp. 35416-35421.
- Ruder, S., 2017. *An overview of gradient descent optimization algorithms*. [Online] Available at: <https://arxiv.org/pdf/1609.04747.pdf> [Acesso em 18/12/2020].
- Saúde, M. d., 2020. *O que é COVID-19*. [Online] Available at: <https://coronavirus.saude.gov.br/index.php/sobre-a-doenca> [Acesso em 18/12/2020].
- Several, 2020. *Sensitivity and specificity*. [Online] Available at: https://en.wikipedia.org/wiki/Sensitivity_and_specificity [Acesso em 18/12/2020].
- Silva, K. R. d. et al., 2020. Alocação de recursos para assistência à saúde em tempos da pandemia de COVID-19: revisão integrativa. *Revista Brasileira de Enfermagem*, Volume 73(Suppl. 2).
- Srivastava, N. et al., 2014. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(56), pp. 1929-1958.
- Sutskever, I., Martens, J. D. G. & Hinton, G., 2013. *On the importance of initialization and momentum in deep learning*. Atlanta, s.n.
- Trujillano, J., March, J. & Sorribas, A., 2004. Methodological approach to the use of artificial neural networks for predicting results in medicine. *Medicina Clinica*, Suppl 1(122), pp. 59-67.