



Previsão de desfecho para pacientes internados com covid-19

Predicting outcomes for patients hospitalized with COVID-19

Predicción de resultados en pacientes hospitalizados por COVID-19

Vitoria Martins Rios¹, Matheus Figueiredo Nunes de Carvalho², Rafael Dutra Ramos³,
Thiago Medeiros Carvalho⁴, Cristiane Oliveira Faria⁵, Karla Tereza Figueiredo Leite⁵

1 Undergraduate Student, Mathematical and Statistics Institute, State University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil

2 M.D., Posto de Saúde Municipal de Maricá, Rio de Janeiro (RJ), Brazil

3 M.D., Hospital Federal da Lagoa, Rio de Janeiro (RJ), Brazil.

4 PhD Student/Lecturer, Mathematical and Statistics Institute, State University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil

5 PhD/Associate Professor, Mathematical and Statistics Institute, State University of Rio de Janeiro, Rio de Janeiro (RJ), Brazil.

Autor correspondente: PhD. Karla Tereza Figueiredo Leite

E-mail: karlafigueiredo@ime.uerj.br

Resumo

Objetivo: Este estudo tem como objetivo avaliar a eficácia de modelos de *Machine Learning* (ML) na previsão de desfechos para pacientes diagnosticados com COVID-19 considerando dados de prontuários e exames. Método: Foram investigados diversos algoritmos de ML e técnicas *Explainable* em dados de evolução clínica dos pacientes internados no Hospital Universitário Pedro Ernesto (HUPE) durante os anos de 2020 e 2021. Resultados: O modelo *Random Forest* foi considerado o mais eficiente, com uma acurácia de 74% na etapa de validação. Além disso, técnicas baseadas em *Explainable Artificial Intelligence* mostram que alterações no número de bastões e a prescrição de noradrenalina foram as variáveis que tiveram maior impacto na previsão de desfechos. Conclusão: Os resultados incentivam instituições de saúde a usarem métodos baseados em apoio à decisão para organizarem ou mesmo priorizarem o atendimento de seus pacientes.

Descritores: Machine Learning; Desfecho COVID-19; Explainable

Abstract



Objective: This study aims to evaluate the effectiveness of Machine Learning (ML) models in predicting outcomes for patients diagnosed with COVID-19 considering data from medical records and exams. **Method:** Several ML algorithms and Explainable techniques were investigated on clinical evolution data of patients admitted to the Pedro Ernesto University Hospital (HUPE) during the years 2020 and 2021. **Results:** The Random Forest model was found to be the most efficient, with an accuracy of 74% in the validation dataset. In addition, techniques based on Explainable Artificial Intelligence show that changes in the number of rods and the prescription of noradrenaline were the variables that had the greatest impact on predicting outcomes. **Conclusion:** The results encourage healthcare institutions to use methods based on decision support to organize or even prioritize care for their patients

Keywords: Machine Learning; COVID-19 Outcome; Explainable

Resumen

Objetivo: Este estudio tiene como objetivo evaluar la eficacia de los modelos de *Machine Learning* (ML) en la predicción de resultados para pacientes diagnosticados de COVID-19 considerando datos de historias clínicas y pruebas. **Método:** Se investigaron diversos algoritmos de ML y técnicas explicables sobre datos de evolución clínica de pacientes ingresados en el Hospital Universitario Pedro Ernesto (HUPE) durante 2020 y 2021. **Resultados:** El modelo *Random Forest* resultó ser el más eficiente, con una precisión del 74% en la etapa de validación. Además, las técnicas basadas en Inteligencia Artificial Explicable muestran que los cambios en el número de barras y la prescripción de noradrenalina fueron las variables que mayor impacto tuvieron en la predicción de resultados. **Conclusión:** Los resultados animan a las instituciones sanitarias a utilizar métodos basados en el apoyo a la toma de decisiones para organizar o incluso priorizar la atención a sus pacientes.

Descriptores: Machine Learning; Resultado COVID-19; Explainable

INTRODUÇÃO

A pandemia de COVID-19 tem afetado o mundo de maneira significativa desde o início de seu surto no final de 2019. Com mais de 700 milhões de casos registrados e mais de 6 milhões de mortes em todo o mundo, a pandemia impactou a saúde pública, a economia global e a vida das pessoas. Apesar da Organização Mundial da



Saúde (OMS) já ter declarado o final da pandemia, de 1º de janeiro a 16 de março de 2024 foram notificados 483.357 casos de COVID-19 e 2.328 óbitos. Além disso, percebe-se que ainda perduram casos com síndromes pós-covid. Um dos aspectos importantes do gerenciamento de crise de uma pandemia é prever a evolução da doença, identificar pacientes em risco de evolução para óbito e os fatores que levam a esse desfecho. Tal entendimento pode ajudar os profissionais de saúde a se prepararem para uma possível escalada da doença e a alocar recursos materiais e humanos de maneira mais eficiente.

Com a crescente digitalização dos registros médicos, o uso de *Machine Learning* (ML) nessa área tem se expandido significativamente, e os algoritmos de previsão baseados em ML podem, aprendendo com dados de pacientes anteriores, inferir a evolução do estado de saúde de novos pacientes, por exemplo.

Fundamentação Teórica

Generalized Additive Model

O GAM é uma técnica simples que permite analisar as relações parciais entre a variável de saída, ou seja, a resposta, e as características (atributos de entrada), enquanto controla a precisão das variáveis preditoras. O método consiste em buscar valores ótimos para os coeficientes e para as funções de base, permitindo obter-se uma relação mais clara e compreensível entre as diversas variáveis⁽¹⁾.

Random Forest

O *Random Forest* é um algoritmo de ML supervisionado que constrói, de forma indutiva, várias árvores de decisão (*ensemble*) para serem usadas em problemas de regressão e classificação, construídas independentemente (paralelo). A ideia por trás do uso de um conjunto de árvores é que a combinação de várias delas, usando a técnica de *bagging*, que consiste em amostrar (com reposição) múltiplos conjuntos de dados de treinamento, pode ser mais acurada do que qualquer modelo individual⁽²⁾. Para evitar a correlação entre as árvores, o algoritmo seleciona aleatoriamente um número de atributos disponíveis para construir cada uma delas e seus ramos, considerando ainda a métrica para avaliar quais atributos serão de fato empregados no modelo⁽³⁾.



XGBoost

O *XGBoost* é um algoritmo de ML baseado no algoritmo de árvores de decisão, que faz parte da família dos métodos de ensemble. Este método opera em série, criando árvores de decisão que aprimoram modelos anteriores por meio da minimização do erro de gradiente decrescente. Este algoritmo pode incluir a penalização de modelos mais complexos, como LASSO (L1) e *Ridge* (L2) para prevenir o sobreajuste⁽⁴⁾. Algoritmos de *boosting* são preferíveis aos de *bagging* devido à sua capacidade de lidar com problemas de viés e variância nos modelos, além de sua habilidade em lidar com esparsidade, garantindo uma melhor escolha de direção para os nós na árvore.

Multi-Layer Perceptron

O *Multi-Layer Perceptron* (MLP) é uma Rede Neural em camadas na qual a informação flui da entrada para a camada de saída, passando por camadas ocultas. Cada conexão entre neurônios possui seu próprio peso, e *perceptrons* na mesma camada possuem a mesma função de ativação. Geralmente, a função de ativação utilizada nas camadas ocultas é a função sigmoide (atualmente também se destaca o uso da função *relu*⁽⁵⁾). Dependendo da aplicação, a camada de saída pode ser uma função sigmoide ou linear⁽⁶⁾.

Entre os muitos algoritmos de aprendizado, o amplamente conhecido método utilizado no treinamento do MLP é o *backpropagation*, que é uma generalização da regra do Mínimo Quadrado Médio. Os pesos são ajustados retropropagando os erros através das camadas da saída para a entrada, daí o nome *backpropagation*⁽⁶⁾.

O desempenho do modelo depende de vários fatores, como a seleção das variáveis, o número de camadas ocultas, a taxa de aprendizado, o momentum e o uso de *early-stopping*, que controla a alteração dos pesos, além do número de iterações.

Shapley Additive Explanations

O método SHAP é utilizado para explicar a importância de cada variável em um modelo, identificando quais destas têm maior impacto no resultado da predição. Ele é particularmente útil em modelos considerados *black-box*, em que é difícil entender como as decisões estão sendo tomadas. Este método é baseado no conceito de teoria



dos jogos e atribui valores a cada recurso, considerando todas as possíveis combinações de recursos para uma dada predição⁽⁷⁾.

O trabalho de Silva (2021), por exemplo, explorou dados disponibilizados pelo Sistema Único de Saúde para prever evolução para óbito ou cura, por algoritmos de ML, utilizando variáveis clínicas e demográficas, alcançando 81% de acurácia com Modelos XGBoost e Redes Neurais⁽⁸⁾.

Na pesquisa proposta por Wollenstein-Betech, et al. (2020) foi obtida uma acurácia de 72% para prever mortalidade com o algoritmo *Support Vector Machine* (SVM), e em torno de 76% de acurácia, para a maioria dos algoritmos usados, para prever o uso de ventilação. Esses resultados foram obtidos ao considerar dados demográficos, comorbidades, sintomas, e informações clínicas (e.g., achados em raio-x)⁽⁹⁾.

No trabalho apresentado por Hu et al. (2020), o objetivo dos autores foi desenvolver um modelo para prever precocemente a evolução de pacientes com infecção grave por COVID-19, baseado em dados demográficos, clínicos e dos primeiros exames laboratoriais após a admissão. As variáveis idade, hsCRP (da sigla em inglês *High-sensitivity C-reactive protein*), d-dímero e contagem de linfócitos se destacaram e foram usadas em um modelo de Regressão Logística, atingindo valores de sensibilidade e especificidade de 83,9% e 79,4%, respectivamente para o conjunto de validação⁽¹⁰⁾.

Este trabalho tem como objetivo comparar o resultado da classificação de risco de óbito em pacientes com COVID-19 usando modelos baseados em ML, considerando informações clínicas obtidas na admissão da internação. Os atributos utilizados estão relacionados ao uso de medicamentos, comorbidades, sinais e sintomas e medicamentos ministrados no período de internação. Além disso, de forma diferenciada dos demais trabalhos, foram exploradas técnicas baseadas em *Explainable Artificial Intelligence* (XAI), visando determinar quais fatores são relevantes para esse resultado.

O restante deste documento está dividido em mais 4 seções. A seção dois faz um resumo dos fundamentos teóricos e apresenta a metodologia utilizada neste trabalho para realizar a avaliação do desfecho para COVID-19. A terceira seção



descreve o Estudo de Caso e, por fim, a última seção conclui e apresenta perspectivas de trabalhos futuros.

MÉTODOS

Foram escolhidos quatro algoritmos para estimar o desfecho da COVID-19 . Apenas um deles, o Modelo Aditivo Generalizado (do inglês, *Generalized Additive Model* - GAM), é um modelo *glass-box* e pode ser naturalmente explicado⁽¹¹⁾. Os outros três modelos são *black-box* e não são naturalmente explicáveis São eles: *Multi-Layer Perceptron* (MLP)⁽⁶⁾, *Random Forest* (RF)⁽³⁾ e *Extreme Gradient Boosting* (XGBoost)⁽⁴⁾.

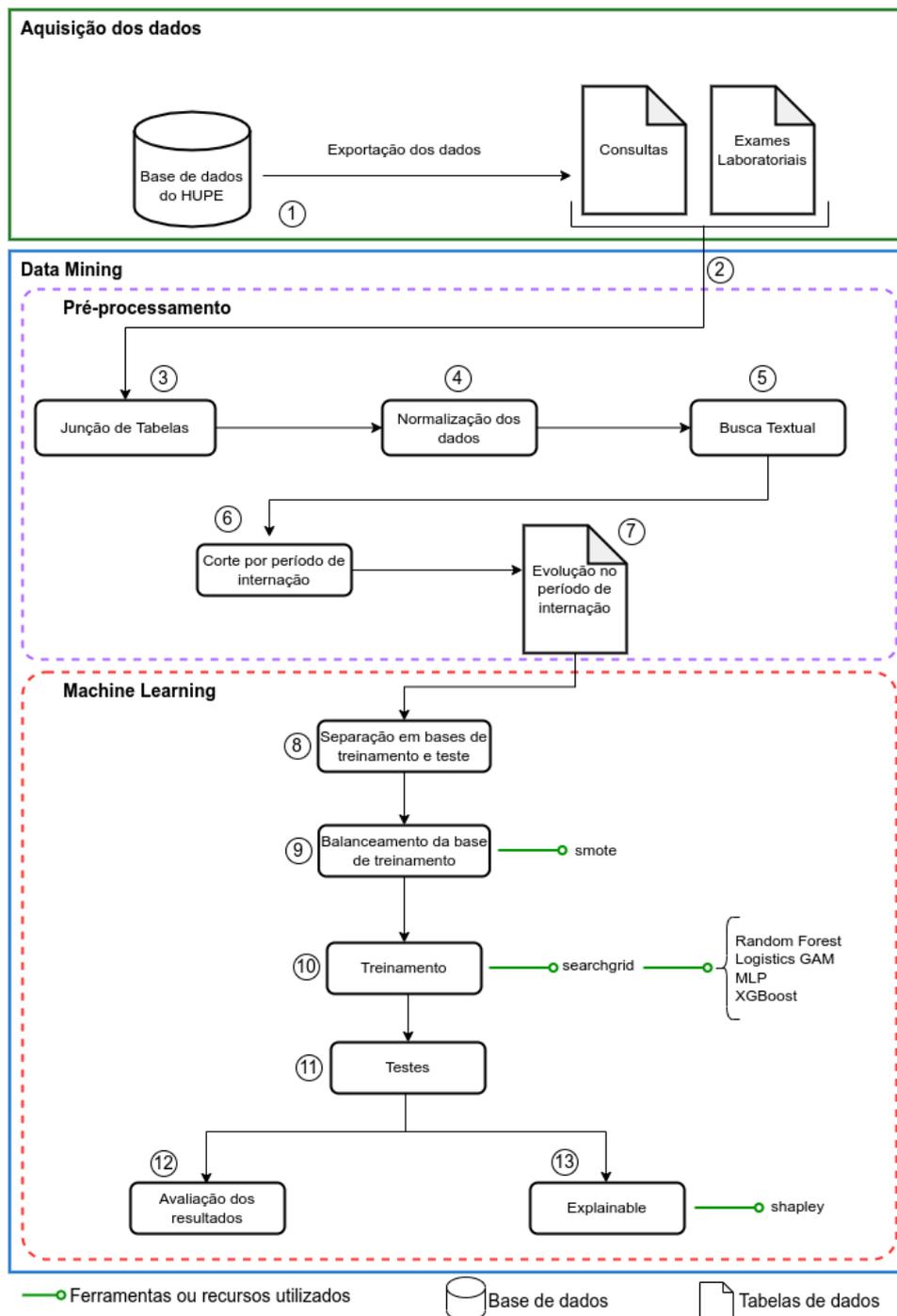
Para alcançar a explicabilidade que está faltando nesses modelos *black-box*, foi utilizado o método *Shapley Additive Explanations* (SHAP)⁽⁷⁾. Além disso, foi utilizada a técnica de validação cruzada (do inglês, *cross validation*) para selecionar os melhores hiperparâmetros de cada algoritmo. Essa combinação de modelos busca explorar modelos lineares e não lineares, além de outros baseados em *ensemble* (RF e XGBoost).

Validação Cruzada

A validação cruzada⁽¹²⁾ é uma técnica utilizada na avaliação de modelos. Neste processo, são feitos testes com diversos hiperparâmetros e, em seguida, é realizada uma avaliação entre todos os modelos para definir o que produz o melhor desempenho. Para avaliar o desempenho, são utilizadas diversas métricas, sendo as mais comuns as médias dos resultados de precisão, recall, acurácia e f1-score⁽⁶⁾.

Em seguida, foi realizada uma busca de melhores hiperparâmetros para os quatro algoritmos de ML utilizados neste trabalho. Nesta abordagem, é possível comparar o desempenho dos diferentes modelos e selecionar aquele que apresenta o melhor resultado para o problema em questão. A Figura 1 apresenta a metodologia adotada no presente estudo, evidenciando as etapas utilizadas neste processo.

Figura 1 – Esquemático da metodologia: etapas da criação dos modelos de ML



Aquisição de dados

A base de dados utilizada nesta pesquisa foi obtida a partir dos registros de pacientes internados no Hospital Universitário Pedro Ernesto (HUPE) entre março de 2020 e agosto de 2021, e que receberam o diagnóstico positivo de COVID-19. Essa



base foi disponibilizada pelo próprio hospital e possui tabelas de dados relativas a pacientes, consultas e exames laboratoriais.

Pré-processamento

Um trabalho de junção dessas tabelas foi realizado a fim de agregar informações acerca de toda a evolução durante esse período de tempo. É importante ressaltar que, para preservar a privacidade dos pacientes, dados de identificação e características pessoais foram removidos deste estudo. Além disso, esta pesquisa foi previamente aprovada pelo Comitê de Ética em Pesquisa da Universidade do Estado do Rio de Janeiro (UERJ) (CAAE: 35192920.2.0000.5282).

Neste primeiro momento, foi observado que as bases de dados possuem informações não padronizadas. Essa falta de padronização dificultou a identificação automática das informações relevantes para a análise, tornando necessária a utilização de uma abordagem manual. Esta etapa consistiu na seleção das informações com base nos mecanismos de busca textual. Essa busca permitiu a extração das informações necessárias (e.g., medicações ministradas, comorbidades) para a construção de uma base consolidada de evoluções para a aplicação dos algoritmos de ML. Além disso, outras informações na base de exames também foram extraídas.

Por fim, foi realizado um recorte temporal na base consolidada de evoluções (prontuário dos pacientes) com o objetivo de selecionar pacientes que tiveram um desfecho após esse período de tempo. Esta escolha foi baseada na quantidade de dados coletados no período e o tamanho da base após o recorte, pois infelizmente a base não tinha o volume de dados adequado para se adotar qualquer período de tempo de internação.

Neste trabalho, a base de dados original contém um desbalanceamento entre os desfechos (óbito/cura), o que pode resultar em um viés para a classe majoritária e prejudicar a acurácia da classificação. Para lidar com esse problema, a base de treinamento foi balanceada utilizando a técnica *Synthetic Minority Over-sampling Technique* (SMOTE)⁽¹³⁾, que consiste na geração de amostras sintéticas baseado em uma combinação linear de vizinhos mais próximos ao dado com atributos ausentes, de forma que as classes fiquem com a mesma quantidade de amostras.



Processamento dos Algoritmos

Após o balanceamento das bases de treinamento, foram realizados ajustes dos modelos baseados em quatro diferentes algoritmos de aprendizado de máquina para a classificação de dados, com o objetivo de avaliar e comparar a eficácia dos algoritmos na classificação dos dados e identificar aquele que possui a melhor solução.

Para encontrar os melhores parâmetros para cada um dos algoritmos, foi utilizado o método *Searchgrid* da biblioteca *Scikit-learn*. Para o algoritmo GAM, foi utilizada a biblioteca *pygam*. Em ambos os casos, o ambiente utilizado continha a versão 3.9.16 do (Python). Esse método envolve a avaliação de diversos modelos a partir da combinação exaustiva de um conjunto de valores para cada hiperparâmetro.

Além disso, para avaliar a importância das variáveis na classificação dos dados, utilizou-se o método SHAP (com base na biblioteca Shap), de forma a avaliar a contribuição de cada variável na predição do modelo.

Após a extração das informações clínicas dos prontuários, foram incluídos atributos que possam auxiliar na classificação do desfecho do paciente, como dados demográficos e sintomas apresentados.

Os exames de imagem não foram considerados no estudo, pois muitos pacientes não o possuíam, o que tornaria inviável a investigação nesse momento, reduzindo ainda mais a base de dados.

A base de dados contém 829 pacientes com diagnóstico de COVID-19 e 62 variáveis de análise diferentes, distribuídas entre comorbidades, medicações prévias, medicações prescritas durante a internação, sintomas e exames laboratoriais.

As informações de medicações prévias, comorbidades e sintomas prévios foram mapeadas na admissão do paciente e inseridas na base de dados com o uso da busca textual. Com base nas consultas subsequentes, foram mapeadas informações sobre a presença de medicamentos ministrados durante a internação. Para os atributos relacionados aos exames laboratoriais, foi considerado apenas o período de internação do paciente avaliado para ser gerar o desfecho. Se houvesse pelo menos um resultado que se encontrasse fora dos valores padrões, a variável correspondente seria mapeada para 1, caso contrário, seria mapeada para 0.

RESULTADOS



Para fins de análise, foi considerado um recorte de 20 dias de internação de forma que apenas pacientes que tiveram desfecho após esse período fossem analisados, sendo 386 dias o máximo período de tempo de internação observado. Este período de tempo foi escolhido por abranger uma quantidade considerável de pacientes em função da quantidade de variáveis que poderiam ser coletadas.

Após o recorte, a base resultante consistiu em 322 pacientes, dos quais 119 vieram a óbito e 203 foram curados. Essa base foi dividida em um conjunto de treinamento e validação (validação cruzada), sendo 80% destinados ao conjunto de treinamento e 20% (66 pacientes - 33 óbitos e 33 curados) para validação usando dados reais.

Além disso, para evitar viés nos resultados devido à desigualdade na distribuição de pacientes em cada classe foi utilizado o método SMOTE para balancear apenas a base de treinamento (óbito), ou seja, a base de validação não foi modificada. Após os balanceamentos da classe minoritária, as bases de treinamentos consistiram em 170 pacientes de cada classe, totalizando 340 desfechos (cura e óbito).

Para a avaliação dos resultados, usando diferentes modelos, foram utilizadas as métricas de precisão, *recall*, *F1-score*, acurácia e *Area Under the Receiver Operating Characteristic Curve* (AUROC). Estas medidas permitem avaliar diferentes aspectos do desempenho do modelo e fornecem uma visão abrangente de sua eficácia⁽¹⁴⁾.

As Tabelas numeradas de 1 a 4 apresentam a relação de hiperparâmetros avaliados para os modelos *Random Forest*, GAM, MLP e *XGBoost* respectivamente. A tabela 5 apresenta os melhores resultados a partir do valor médio obtido com validação cruzada para cada um dos quatro modelos. O modelo de *Random Forest* obteve o melhor desempenho em relação aos demais modelos, tendo se sobressaído em todas as métricas avaliadas. Por outro lado, o modelo GAM apresentou um desempenho inferior aos demais, com uma acurácia de aproximadamente 69%. Esses resultados indicam que o algoritmo de *Random Forest* é o mais adequado para prever os desfechos clínicos em pacientes hospitalizados. Porém, os modelos de *XGBoost* e MLP também apresentaram desempenhos semelhantes e podem ser considerados alternativas viáveis.

Tabela 1 – Avaliação de hiperparâmetros para o modelo *Random Forest*



Parâmetro	Função	Valores testados	Melhor valor
max depth	Máxima profundidade de cada árvore	6, 8, 10, 15, 20, 25, 30, 35, 40	25
min samples split	Número mínimo de amostras requerido para subdividir um nó	1, 2, 4, 8	2
n estimators	Número de árvores de decisão	25, 50, 100, 200	50
criterion	Função utilizada para medir a qualidade de uma divisão	gini, entropy	entropy
max features	Número máximo de rótulos considerados	2, 4, 6, 8, 10	2
min samples leaf	Número mínimo de amostras necessário para definir um nó	1, 2, 4	1

Tabela 2 – Avaliação de hiperparâmetros para o modelo GAM

Parâmetro	Função	Valores testados	Melhor valor
max iter	Número máximo de iterações permitidas para que o modelo convirja	1, 10, 50, 100, 200, 300	10
tol	Tolerância para os critérios de parada.	1, 10, 50, 100, 300, 500	1

Tabela 3 – Avaliação de hiperparâmetros para o modelo MLP

Parâmetro	Função	Valores testados	Melhor valor
hidden layer sizes	Tamanho da camada escondida	(50,50,50), (50, 100, 50, 100), (50,100,50), (100, 100, 100), (100,)	(50, 100, 50)
activation	Função de ativação da camada escondida	relu, tahn	relu
solver	Solucionador para otimização de pesos	adam, sgd, lbfgs	adam
alpha	Força do termo de regularização L2	0,0001, 0,001, 0,01, 0,1	0.001
learning rate	Definição da taxa de aprendizagem para atualizações de pesos	constant, adaptative	constant
nesterovs momentum	Utilização de Nesterov's momentum, somente se o solver for 'sgd'	True, False	-

Tabela 4 – Avaliação de hiperparâmetros para o modelo XGBoost



Parâmetro	Função	Valores	
		testados	Melhor valor
base score	A pontuação de predição inicial de todas as instâncias, viés global	0,25, 0,5, 0,75, 1,0	0.75
min child weight	Soma mínima do peso da instância necessária em um nó filho	1, 3	1
subsample	Proporção de subamostragem das instâncias de treinamento	0,5, 0,75, 1,0	1.0
colsample bytree	Proporção de subamostragem de colunas ao construir cada árvore	0,5, 0,75, 1,0	0.75
n estimators	Número de árvores a serem ajustadas na floresta	100, 200, 300, 400	350
learning rate	Taxa de aprendizado do algoritmo	0,1, 0,001, 0,0001	0.1
max depth	Profundidade máxima da árvore	5, 10, 15, 20, 30	15

Tabela 5 – Resultados da avaliação dos modelos

Modelo	Precisão	F1	Recall	Acurácia	ROC AUC
RF	0,7305	0,7510	0,7828	0,7452	0,7999
GAM	0,6968	0,6894	0,6894	0,6894	0,6894
MLP	0,7074	0,6964	0,7136	0,7141	0,7427
XGBoost	0,7087	0,7224	0,7458	0,7205	0,7507

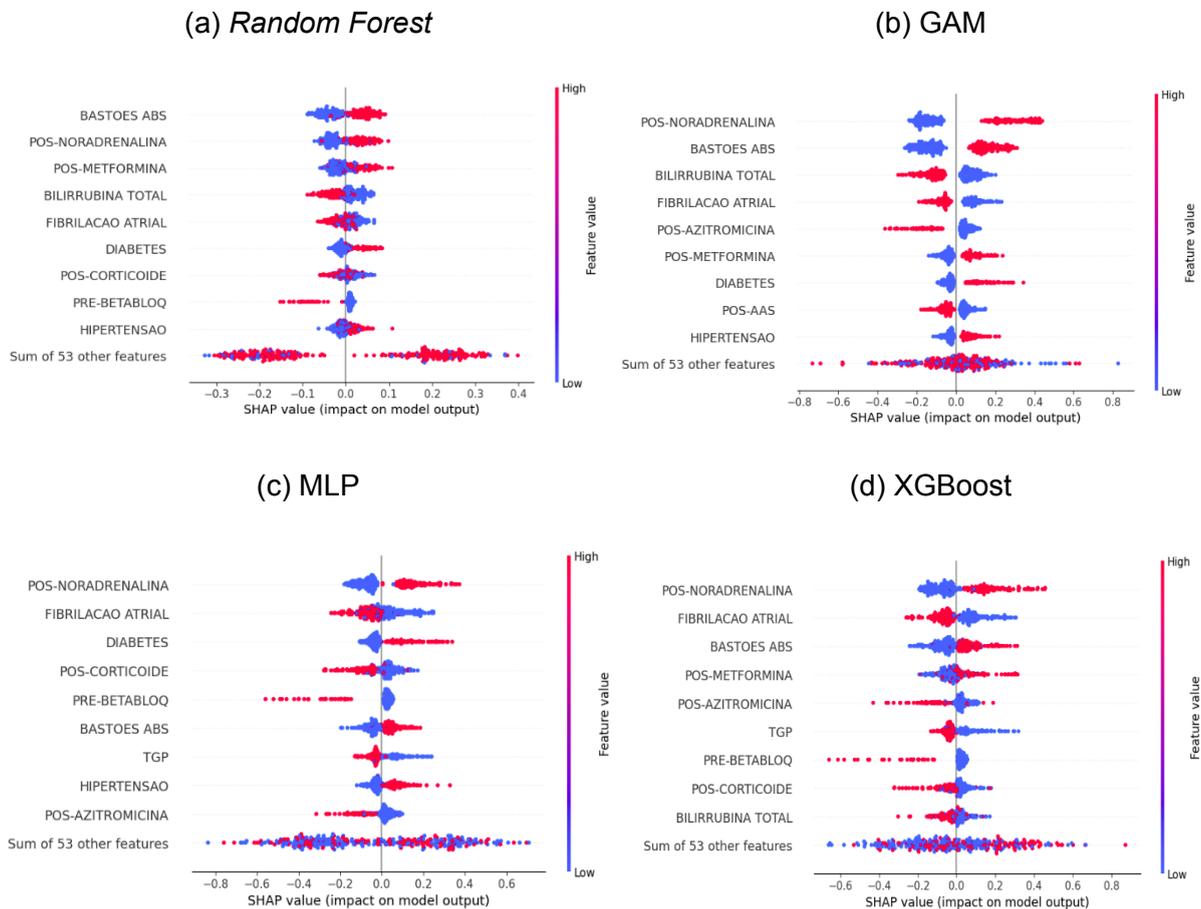
Neste trabalho, foi realizada uma avaliação do impacto das variáveis utilizadas na previsão dos desfechos clínicos para cada modelo. Dessa forma, pôde-se observar que algumas variáveis apresentaram maior influência no resultado da previsão do que outras. A Figura 2 (a) indica as variáveis que melhor explicam os resultados obtidos. No entanto, a título de avaliação, a Figura 2 apresentam essa avaliação também para os modelos GAM, MLB e *XGBoost*, respectivamente.

Entre as variáveis que tiveram maior impacto na previsão, destacam-se aquelas que indicam a presença de doenças prévias ou ocorrências observadas na internação, como diabetes e fibrilação atrial, além dos medicamentos prescritos durante a internação, como noradrenalina, normalmente prescrita na modulação da pressão arterial, metformina (normalmente indicado para o tratamento da diabetes mellitus tipo 1 ou tipo 2, além da observação do número de bastões nos exames de laboratório. Essas variáveis são consideradas importantes por representarem fatores de risco para complicações durante a internação. Além disso, também foi observado que as



alterações nos valores dos exames de bilirrubina total e bastões tiveram grande impacto na previsão.

Figura 2 – Impacto de variáveis no modelo



CONCLUSÃO

A avaliação dos modelos permitiu a seleção do *Random Forest* como o algoritmo mais eficiente para classificação de pacientes com COVID-19 e obter resultados promissores. Acrescenta-se a isso a identificação dos atributos que mais contribuíram para explicar os resultados obtidos, em cada um dos modelos, envolvendo medicações ministradas durante a internação, ocorrências ou resultados obtidos em exames de laboratório. Tornando mais explicável os resultados obtidos, aspecto atualmente considerado fundamental principalmente nos modelos aplicados à saúde. Além disso, os resultados podem incentivar instituições de saúde a usarem métodos baseados em apoio à decisão para organizarem ou mesmo priorizarem



atendimento de seus pacientes. Destaca-se nesse ponto o diferencial em relação aos outros trabalhos já publicados que não consideram medicamentos usados antes ou durante as internações pelos pacientes, além de muitas ocorrências observadas nos prontuários.

No entanto, é importante ressaltar que a baixa quantidade de pacientes analisados, falta de padronização na base de dados original são fatores que podem ter contribuído para esses resultados, limitando a generalização do modelo.

Esses resultados evidenciam a necessidade de aprimoramento das técnicas de coleta e padronização dos dados, o que pode ser realizado através de um estudo mais refinado de coleta e configuração dos parâmetros dos modelos. O uso de seleção de variáveis, que consiste em identificar as variáveis mais importantes para a predição e descartar aquelas que possuem pouca influência no resultado, podem ajudar a melhorar o desempenho do modelo⁽¹⁵⁾.

Dessa forma, trabalhos futuros podem explorar essas possibilidades, além de desenvolver e usar outras técnicas⁽¹⁶⁾ que permitam melhores resultados na classificação de desfechos de pacientes com COVID-19 ou qualquer outra doença de interesse, considerando a característica da *explainability* e, portanto, maior transparência e compreensão sobre os resultados.

Agradecimentos

O presente trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq 308717/2020-1, da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e CAPES 88881.506840/2020-01.

Referências

1. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. In Springer, chapter 15; 2009.
2. Quinlan J R. Induction of decision trees. *Machine Learning*. 1986;1(1):81–106.
3. Breiman L. Random Forests. *Machine Learning*. 2001; 45 (1):5–32.
4. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'16, New York, NY, USA. Association for Computing Machinery*. 2016; 785–794.



5. Fukushima K, Miyake S. Neocognitron: A Self-Organizing Neural Network Model for a Mechanism of Visual Pattern Recognition. Competition and Cooperation in Neural Nets. Lecture Notes in Biomathematics. v 45. Springer. 1986;267–285.
6. Haykin S. Redes Neurais. 2ed. Bookman; 2001.
7. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, Advances in Neural Information Processing Systems. Curran Associates, Inc.; v.30; 2017.
8. Silva AO, Santos BS, Tondato R, Lima RHP. Uso de machine learning para previsão da evolução de casos de srag incluindo casos de COVID-19 considerando variáveis clínicas e demográficas. Trabalho de conclusão de curso. Universidade Tecnológica Federal do Paraná (RIUT); 2021.
9. Wollenstein-Betech S, Silva AAB, Fleck JL, Cassandras CG, Paschalidis IC. Physiological and socioeconomic characteristics predict COVID-19 mortality and resource utilization in Brazil. PLOS ONE; 2020;15(10):e0240346.
10. Hu C, Liu Z, Jiang Y, Shi O, Zhang X, Xu K, Suo C, Wang Q, Song Y, Yu K, Mao X, Wu X, Wu M, Shi T, Jiang W, Mu L, Tully DC, Xu L, Jin L, Li S, Tao X, Zhang T, Chen X. Early prediction of mortality risk among patients with severe COVID-19, using machine learning. International Journal of Epidemiology. 2020; 49(6):1918–1929.
11. Hastie T, Tibshirani R. Generalized Additive Models. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. Taylor & Francis; 1990.
12. Stone M. Cross-Validatory Choice and Assessment of Statistical Predictions". Journal of the Royal Statistical Society, Series B (Methodological). 1974; 36 (2): 111–147.
13. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research. 2002;16: 321–357.
14. Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In Losada, D. E. and Fernández-Luna, J. M., editors, Advances in Information Retrieval. Springer Berlin Heidelberg. 2005; 345–359.
15. Liu H, Motoda, H. Perspectives of Feature Selection. Springer US, Boston, MA. 1998; 17-41.
16. Holanda WD, Silva LC, César Sobrinho AAC. Estratégias Preditivas na Detecção do Agravamento do Quadro Clínico de Pacientes com COVID-19: Uma Revisão de Escopo J. Health Inform. 2021 Outubro-Dezembro; 13(4): 128-32.