



Universidade do Estado do Rio de Janeiro

Centro Biomédico

Faculdade de Ciências Médicas

Erito Marques de Souza Filho

**Desenvolvimento de uma plataforma digital como apoio à triagem de
pacientes com sinais e sintomas sugestivos de infecção pelo Coronavírus
Covid-19 utilizando Machine Learning**

Rio de Janeiro

2022

Erito Marques de Souza Filho

Desenvolvimento de uma plataforma digital como apoio à triagem de pacientes com sinais e sintomas sugestivos de infecção pelo Coronavírus Covid-19 utilizando Machine Learning

UERJ

Orientadora: Prof^ª. Dra. Alexandra Monteiro

Rio de Janeiro

2022

CATALOGAÇÃO NA FONTE
UERJ/REDE SIRIUS/BIBLIOTECA CB-A

S729 Souza Filho, Erito Marques de.

Desenvolvimento de uma plataforma digital como apoio à triagem de pacientes com sinais e sintomas sugestivos de infecção pelo Coronavírus Covid-19 utilizando Machine Learning / Erito Marques de Souza Filho – 2022.

35 f.

Orientadora: Profª. Dra. Alexandra Monteiro

Dissertação (Mestrado) – Universidade do Estado do Rio de Janeiro, Faculdade de Ciências Médicas. Pós-graduação em Telemedicina e Telessaúde.

1. COVID-19 (doenças) – Teses. 2. Sistemas de Informação em Saúde. 3. Aprendizado de máquina - Teses. I. Monteiro, Alexandra. II. Universidade do Estado do Rio de Janeiro. Faculdade de Ciências Médicas. III. Título.

CDU 616-071:004.8

Bibliotecária: Felipe Caldonazzo
CRB7/7341

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta dissertação, desde que citada a fonte.

Assinatura

Data

Erito Marques de Souza Filho

Desenvolvimento de uma plataforma digital como apoio à triagem de pacientes com sinais e sintomas sugestivos de infecção pelo Coronavírus Covid-19 utilizando Machine Learning

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre Profissional em Telemedicina e Telessaúde, ao Programa de Pós-graduação em Telemedicina e Telessaúde, da Universidade do Estado do Rio de Janeiro.

Aprovada em 24 de junho de 2022.

Banca Examinadora

Prof^a. Dra. Alexandra Monteiro
Faculdade de Ciências Médicas - UERJ

Prof^a. Dra. Rosa Maria Esteves Moreira da Costa
Instituto de Matemática e Estatística - UERJ

Prof. Dr. João Carlos Tress
Complexo Hospitalar de Niterói

Rio de Janeiro

2022

DEDICATÓRIA

Dedico esse trabalho a todos aqueles que contribuíram direta ou indiretamente para sua execução.

AGRADECIMENTOS

A todos que contribuíram direta ou indiretamente para a conclusão desse trabalho.

A melhor forma de prever o futuro é criá-lo.

Peter Drucker

RESUMO

SOUZA FILHO, Erito Marques de. *Desenvolvimento de uma plataforma digital como apoio à triagem de pacientes com sinais e sintomas sugestivos de infecção pelo Coronavírus Covid-19 utilizando Machine Learning*. 2022. 35 f. Dissertação (Mestrado Profissional em Telemedicina e Telessaúde) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

A pandemia provocada pelo Covid-19 elencou novos desafios emergenciais sócio-econômicos assim como nas aplicações e uso das tecnologias em saúde. Nesse contexto, a identificação precoce de sinais e sintomas sugestivos desta doença pode favorecer de maneira significativa o processo de triagem e assim contribuir para a prevenção do contágio assim como para o diagnóstico precoce e redução de desfechos clínicos indesejados. O uso de ferramentas de *Machine Learning* (ML) tem tido diversas aplicações na Medicina, em particular em relação ao diagnóstico e prognóstico de pacientes e para a automação de processos como a triagem por algoritmos inteligentes. Nesse trabalho foi desenvolvido um modelo de ML para avaliar a probabilidade do diagnóstico de COVID-19 a partir de sinais e sintomas de pacientes.

Palavras-chave: Aprendizado de máquina. Programas de Triagem Diagnóstica. Infecção por Coronavírus 2019-nCoV. Sinais e Sintomas.

ABSTRACT

SOUZA FILHO, Erito Marques de. *Development of a digital platform to support the screening of patients with signs and symptoms suggestive of Covid-19 infection using Machine Learning*. 2022. 35 f. Dissertação (Mestrado Profissional em Telemedicina e Telessaúde) – Faculdade de Ciências Médicas, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2022.

The Covid-19 pandemic highlighted new socio-economic emergency challenges as well as in the applications and use of health technologies. In this context, early identification of signs and symptoms suggestive of this disease can significantly favor the screening process and thus contribute to the prevention of contagion and the early diagnosis and reduction of unwanted clinical outcomes. The use of Machine Learning (ML) tools has several applications in medicine concerning the diagnosis and prognosis of patients and for the automation of processes such as screening by intelligent algorithms. We proposed a ML model to assess the probability of the diagnosis of COVID-19 from the signs and symptoms of patients.

Keywords: Machine Learning; Diagnostic Screening Programs; COVID-19; Signs and Symptoms

LISTA DE ILUSTRAÇÕES

Tabela 1 –	Características de ferramentas utilizadas para <i>screening</i> de COVID-19	18
Figura 1 –	Tela de apresentação do protótipo	24
Figura 2 –	Tela de identificação e marcação dos sintomas	25
Figura 3 –	Tela de solicitação de outras informações e do aceite das políticas de privacidade	25
Figura 4 –	Tela de saída com a estimativa de probabilidade	26

LISTA DE ABREVIATURAS E SIGLAS

AUC	Área abaixo da curva ROC
AVE	Acidente Vascular Encefálico
DL	<i>Deep Learning</i>
GA	Google Analytics
IA	Inteligência Artificial
INPI	Instituto Nacional de Propriedade Industrial
ML	Machine Learning
SSL	Secure Socket Layer
UPG	Unidades de Processamento Gráfico

SUMÁRIO

	INTRODUÇÃO	11
1	REVISÃO DE LITERATURA	14
2	OBJETIVOS	20
2.1	Geral	20
2.2	Específicos	20
3	MATERIAL E MÉTODOS	26
4	RESULTADOS	24
5	DISCUSSÃO	27
	CONCLUSÕES	29
	REFERÊNCIAS	30

INTRODUÇÃO

Em dezembro de 2019, um grupo de pessoas na cidade de Wuhan, na China, manifestou sinais e sintomas de acometimento respiratório de curso clínico grave e desfecho avassalador que foi diagnosticado inicialmente como pneumonia por causa desconhecida e posteriormente pelo Coronavírus Covid-19 (Zhu et al, 2020). A disseminação foi rápida na China e no mundo, recebendo posteriormente o *status* de pandemia pela Organização Mundial de Saúde. O agente causador foi identificado como pertencente à família dos coronavírus, cuja apresentação clínica guarda relação com aquela observada durante a epidemia de síndrome respiratória aguda grave em 2003 (Guan et al, 2020). A nova enfermidade foi inicialmente chamada de SARS-COV-2 e renomeada posteriormente para COVID-19 pela Organização Mundial de Saúde. Trata-se de uma doença infecciosa com expressão clínica variável que tem a febre e a tosse (geralmente seca) como sintomas predominantes, mas que também pode cursar com diarreia, náuseas, vômitos, dispneia, fadiga, mialgia, anosmia/hiposmia, ageusia/hipogeusia. (Richardson et al, 2020; Wang et al, 2020; Guan et al, 2020).

O Ministério da Saúde definiu em suas “Diretrizes para diagnósticos e tratamento da COVID-19”, publicadas em 2020, que os casos suspeitos da doença seriam àqueles de indivíduos com: a) síndrome gripal: entendida como um quadro respiratório agudo com sensação febril ou febre e que pode vir acompanhado de tosse, ou coriza ou dor de garganta ou dificuldade respiratória; b) indivíduos com Síndrome Respiratória Aguda Grave, podendo apresentar dispneia, desconforto respiratório e até diminuição da saturação de oxigênio. Essa instituição também destaca que outros sintomas são relacionados a doença, tais como a presença de mialgias, a perda ou diminuição do olfato (anosmia/hiposmia), a presença de distúrbios gastrointestinais (diarreia/náuseas/vômitos) ou a perda ou diminuição do paladar (ageusia/hipogeusia). Em idosos, deve-se considerar ainda sintomas como, síncope, sonolência em excesso, confusão mental, falta de apetite e irritabilidade. O diagnóstico definitivo pode ser obtido por critérios clínico e/ou epidemiológicos e por exames laboratoriais, tais como, biologia molecular, ensaio imunoenzimático, imunocromatografia e imunoenensaio por eletroquimioluminescência (Ministério da Saúde, 2020).

A situação de emergência de saúde pública que se estabeleceu e os novos desafios impostos pela COVID-19 configuram um panorama complexo que demanda um conjunto de ações coordenadas e possui a inovação como um de seus pilares. Nesse contexto, as

ferramentas de Inteligência Artificial (IA) podem ser de fundamental valia. Esta engloba um complexo arcabouço de sofisticados modelos matemático-computacionais que permitem a construção de algoritmos a serem utilizados para emular a realização de diversas tarefas humanas, como por exemplo, o reconhecimento de padrões, a resolução de problemas e a compressão da linguagem. Existe uma subárea da IA chamada de *Machine Learning* (ML), que traz a possibilidade de aprendizado a partir da experiência obtida com bases de dados de grande porte coletadas e processadas de maneira adequada (Souza Filho et al., 2019). Diversas ferramentas de ML foram construídas ao longo do tempo e em particular, vale destacar algumas delas, como o *Support Vector Machine*, *Gradient Boosting*, *Random Forests* e as redes neurais artificiais (Cortes e Vapnik, 1995; Breiman, 2001; Friedmann, 1999; Rumelhart et al., 1986). Além disso, há um grande crescimento desta última, principalmente de uma subclasse de modelos chamada *Deep Learning* (DL) (LeCun, Bengio e Hinton, 2015). Vale destacar vantagens dos modelos de ML: a) escalabilidade e flexibilidade quando comparadas com os métodos bioestatísticos convencionais, o que seguramente traz um grande potencial para realização de tarefas como a estratificação de risco e a avaliação diagnóstica e prognóstica. b) habilidade de analisar diversos tipos de dados, como por exemplo, dados clínicos, ômicos, laboratoriais, dados de imagem e dados demográficos e de incorporá-los em predições de risco e no desenvolvimento de condutas médicas (Ngiam e Khor, 2019). Nesse ínterim, vale destacar que os modelos de ML têm se mostrado bastante úteis e relevantes em diversos cenários. Em particular, com relação à triagem diagnóstica pode-se destacar algumas aplicações importantes.

Than et al. (2019) desenvolveram um algoritmo de ML visando estimar a probabilidade de um determinado paciente apresentar infarto agudo do miocárdio. Para alcançar tal objetivo, o modelo analisa variáveis clínicas e laboratoriais, tais como, idade, sexo, concentração de troponina e taxa de variação da troponina. Trata-se de um estudo multicêntrico com a participação de 11000 pacientes. Foi utilizado um modelo chamando *Gradient Boosting*. No tocante ao desempenho, vale ressaltar que os resultados foram bastante animadores. A performance verificada foi melhor que a obtida por escores tradicionais como o da Sociedade Européia de Cardiologia (ESC 0/3-hour pathway). A sensibilidade obtida foi de 82,5% [74.5-88.8%] enquanto a especificidade foi de 92,2% [90.7-93.5%] (Than et al., 2019).

Hair et al. (2019) em seu trabalho combinaram dois modelos de ML (*Stochastic self-organizing maps* e *Random Forests*) para analisar o perfil clínico de 523 casos de dengue. A ideia era identificar características úteis que poderiam ser utilizadas para desenvolver um

critério de gravidade nesses pacientes. Trata-se de um estudo unicêntrico e retrospectivo. Foram encontrados 4 grupos, com distintos perfis em relação à doença. Dois deles estão relacionados ao surgimento de sintomas leves, outro associado a evolução intermediária e um relacionado a maior gravidade. Eles concluíram que a idade dos pacientes foi considerada pelo algoritmo como uma variável chave. Além disso, reiteraram que os sinais de alarme devem receber um acompanhamento estrito, principalmente em crianças (Hair et al., 2019). Kwon et al. (2018), por sua vez, fizeram um estudo de coorte retrospectiva com 52131 pacientes de dois hospitais diferentes para avaliar se um modelo de ML seria capaz de prever que pacientes estariam em risco de parada cardíaca. O modelo utilizado foram as redes neurais artificiais (DL). Foram utilizadas apenas 4 variáveis, a saber, pressão sistólica, frequência cardíaca e temperatura corporal. No tocante a predição de parada cardíaca, o algoritmo teve um bom desempenho, a área abaixo da curva ROC (AUC) foi de 0,85 no primeiro hospital e 0,83 no segundo. Os autores também avaliaram o desempenho do algoritmo no tocante a predição de morte de um paciente caso não receba ressuscitação cardiopulmonar. As AUC obtidas nesse caso foram, respectivamente, 0,92 e 0,91. Eles reiteraram que os sistemas tradicionais (*track-and-trigger*) possuem algumas limitações como a baixa sensibilidade e a elevada taxa de alarmes falsos e que, portanto, o algoritmo de DL poderia contribuir de maneira significativa para melhorar esse tipo sistema (Kwon et al, 2018).

Considerando tudo o que foi exposto anteriormente, este projeto tem por objetivo desenvolver um algoritmo para triagem diagnóstica de COVID-19, a partir de dados clínicos dos pacientes, utilizando um modelo de redes neurais artificiais (DL). Dessa forma, a ferramenta foi registrada no Instituto Nacional de Propriedade Industrial (INPI).

1 REVISÃO DA LITERATURA

O termo Inteligência Artificial foi cunhado pela primeira vez em uma conferência realizada na cidade de Dartmouth em 1956, entretanto seu desenvolvimento se confunde em partes com a própria história da computação, tendo tido contribuição fundamental de diversos autores, tais como Alan Turing, John Von Neumann, dentre outros (McCarthy et al., 1955; McCorduck, 2004; Souza Filho, 2019). Em particular, com o advento de uma classe de modelos específicos denominada DL permitiu-se que problemas de elevada complexidade na área de saúde pudessem ser resolvidos. As redes neurais têm suas origens nos trabalhos de McCulloch e Pitts que lançaram as bases para a construção de modelos matemáticos que tivessem alguma proximidade com os neurônios biológicos (McCulloch e Pitts, 1943). De modo superficial, sem aprofundar nos detalhes sob o ponto de vista da matemática, pode-se afirmar que a base do modelo é a utilização de um conjunto de entes matemáticos denominados neurônios artificiais. Eles são dispostos em camadas e possuem conexões entre si para os quais são associados valores numéricos denominados pesos. Em última análise, o algoritmo visa encontrar um conjunto de pesos dito ótimos, considerando as conexões existentes na rede neuronal, um conjunto de entradas e um conjunto de funções matemáticas denominadas funções de ativação. O primeiro desses modelos foi proposto por Rosenblatt e denominado “*Perceptron*” (Rosenblatt, 1958). Entretanto, em 1969, Minsky e Papert provaram que o uso de redes neurais dessa natureza teria uma utilidade limitada, na medida em que seus melhores resultados só ocorreriam em problemas que possuíssem uma característica especial: serem linearmente separáveis (Minsky e Papert, 1969). O desenvolvimento das redes neurais artificiais nas décadas posteriores seguiu certo período de latência sob o ponto de vista da pesquisa, ainda que alguns avanços importantes tenham sido alcançados.

O trabalho de Broomhead e Lowe sobre as chamadas Funções de Base Radial, é um exemplo (Broomhead e Lowe, 1988). Essa situação sofreu uma significativa mudança devido a pelo menos três fatores. O primeiro deles foi o trabalho seminal de Rumelhart e cols, que permitiu a realização da atualização dos pesos em complexas redes com múltiplas camadas de neurônios – DL – denominado *backpropagation* (Rumelhart et al, 1986). Além disso, o uso de computação paralela para realização de cálculos complexos de maneira mais rápida foi outro ponto fundamental. Trata-se das chamadas Unidades de Processamento Gráfico (UPGs). A inspiração nesse caso foi nas placas gráficas que já eram utilizadas pela indústria do

entretenimento para produção dos jogos modernos. Alguns desses jogos, como por exemplo, o Grand Theft Auto e Assassin's Creed emulam personagens com grande realismo. O uso das UPGs como paradigma para os modelos de ML foi fundamental para acelerar a obtenção dos resultados dos modelos matemático-computacionais. O terceiro fator fundamental foi a existência de um ambiente de BIG DATA, com grande importância no processo de treinamento dos modelos de ML (Souza Filho, 2019).

A construção das redes neurais artificiais traz no seu bojo pelo menos dois tipos de aprendizado distintos. No primeiro, denominado aprendizado supervisionado, existe um objetivo claro de prever o resultado de um desfecho (ou de uma medida desse desfecho) tendo como entrada uma quantidade de exemplos rotulados. Por exemplo, prever se o indivíduo tem ou não COVID-19 tendo como base os dados obtidos anteriormente relativos a outros indivíduos. O aprendizado supervisionado, por sua vez, tem como objetivo descrever associações e padrões a respeito de um conjunto de dados de entrada, como por exemplo, quando se deseja agrupar um conjunto de indivíduos como relação a um determinado critério, como a propensão a ser infectado por COVID-19 (Hastie et al, 2009). Após a coleta e processamento dos dados a serem utilizados, segue-se uma fase denominada engenharia de atributos. Nessa etapa, definem-se quais variáveis serão utilizadas nos processos de treinamento, validação e teste dos modelos. O processo de treinamento consiste no treinamento dos modelos visando encontrar os melhores ajustes dos seus parâmetros. O processo de validação consiste na avaliação do desempenho dos modelos visando selecionar qual deles teve melhor *performance*. O processo de teste visa avaliar a capacidade de generalização do algoritmo selecionado. No âmbito desse projeto, vale mencionar, que existem trabalhos que não realizam o processo de validação e outros que utilizam técnicas um pouco diferentes, como o processo de validação cruzada (James et al., 2013).

No tocante as ferramentas de acesso gratuito já existentes com relação ao tema, disponíveis em lojas da PlayStore e Apple, pode-se destacar a *Corona CheckApp*. Trata-se de um aplicativo de origem alemã, disponível no iOS, que solicita ao usuário o preenchimento de informações sobre sintomas nas últimas 24 horas (febre, tosse, dispnéia, faringalgia, coriza, perda de olfato ou paladar, cefaleia, dor muscular ou articular ou fadiga), sobre antecedente de contato do usuário com pacientes com a doença bem como idade e sexo. Ao final do preenchimento, o paciente recebe a informação se é considerado ou não suspeito de ter a doença, orientações gerais (distanciamento social e medidas de higiene) e telefones de emergência para contato na Alemanha. O *app* é inglês, mas tem textos em alemão. O C-

Screeener é uma ferramenta de triagem diagnóstica para pacientes com suspeita da doença disponível para Android e iOS. O usuário deve informar seu nome, telefone e email antes de iniciar o uso da ferramenta. Nessa aplicação são solicitadas informações similares as do *Corona Check* sobre a sintomatologia do usuário. Além disso, o aplicativo solicita que o usuário informe nome, email e telefone. Após o preenchimento das informações solicitadas, o usuário recebe via telefone ou mensagem de texto indicando se é considerado suspeito ou não de possuir a doença bem como um conjunto de ações recomendadas (por exemplo, distanciamento social, lavagem das mãos, uso de máscara etc). O *Apple Covid-19*, disponível em português, solicita que o usuário ligue para 911 se tiver apresentando sintomas como dor no peito grave e constante, dificuldade extrema de respirar, tontura constante e grave, desorientação grave, irresponsividade ou rosto e lábios azulados. Caso não presente nenhum desses sintomas, o usuário deve informar a idade, se fez ou não teste para COVID recentemente (e o resultado positivo ou negativo) e quais sintomas se aplicam (febre ou calafrio; dificuldade respiratória leve ou moderada; tosse recente ou agravada; perda repentina de olfato ou paladar; dor de garganta; vômitos ou diarreia; fadiga intensa e inexplicável ou dores por todo corpo). Em seguida, o usuário deve relatar sobre sua percepção se os sintomas tem pouco/nenhum impacto, algum impacto ou grande impacto em sua vida cotidiana. O paciente ainda deve inserir no *app* e tem alguma das seguintes condições: obesidade, tabagismo, gravidez, diabetes, pressão alta, doença renal crônica ou doença hepática, doença pulmonar crônica, condição cardíaca grave, vírus da imunodeficiência humana, tratamento contra câncer, doença cérebro vascular, condição neurológica, anemia falciforme ou outras doenças hematológicas. Na sequência, deve ser informado se houve alguma possibilidade de exposição a doença e se mora em instituição de cuidados prolongados ou se trabalhou ou foi voluntário em instituição de saúde. Ao final o usuário, recebe um conjunto de recomendações, como distanciamento social, monitoramento de sintomas, uso de mascaras, lavagem das mãos, repouso, ingestão hídrica adequada e se é recomendado fazer teste para COVID-19. O aplicativo também permite, de modo opcional, que o usuário informe a sua localização de origem no território nacional. Além disso, traz informações sobre o vírus e um resumo atualizado da doença, dicas gerais para manter a saúde física e mental e permite ao usuário a configuração de um prontuário, dentro do próprio app, com suas informações médicas, preenchido pelo usuário para uso por socorristas em caso de emergência. O *Beebe HealthCare* disponibiliza em inglês um *Chatbot*, que é algoritmo capaz de simular o processo de conversação entre indivíduos, para auxiliar o usuário no rastreamento da doença, que

pergunta ao usuário se está sentindo sintomas que indicam gravidade, similar ao aplicativo da Apple, entretanto com mais de sintomas, como alterações no discurso, sinus e sintomas relativos a hipotensão e tosse com sangue. Caso o paciente apresente algum desses sintomas, o paciente é orientado a ligar para 911. Caso contrário, o paciente é perguntado a respeito da sua idade, outros sintomas (dispnéia, desidratação, febre, dor muscular, perda do odor e paladar, diarreia, náusea e vômitos). Caso positivo, para alguma dessas situações, o paciente é orientado a realizar o teste para COVID-19. O aplicativo possui uma seção contendo informações sobre locais para realização do teste. O *Centers for Disease Control and Prevention* (CDC) possui uma ferramenta chamada *Coronavirus Self Checker*. Trata-se de um aplicativo para rastreamento da doença. Essa ferramenta solicita que o usuário digite se é estadunidense ou se está nos Estados Unidos no momento sua cidade de origem, estado e país. Na sequência o usuário é perguntado sobre sua idade, sexo, origem latina ou hispânica ou espanhola, raça, sobre sintomas de gravidade, se testou positivo para a doença ou se tem sintomas (febre, tosse, dispnéia, faringite, cefaleia, vômito, diarreia, dor muscular, hiposmia ou ageusia, congestão), se reside ou trabalha em asilos (ou similares) ou se trabalha ou trabalhou como voluntário em unidade de saúde nos últimos 14 dias. Em seguida uma lista de comorbidades é apresentada incluindo doenças pulmonares, doenças cardiovasculares, obesidade, diabetes, hipertensão, doença renal, câncer, HIV, doenças hematológicas, doença cerebrovascular, tabagismo, Síndrome de *Down* e presença de imunossupressão. Ao final o paciente recebe informações gerais sobre a doença e conduta; para pacientes que residam ou estejam nos Estados Unidos, há um direcionamento para o site do estado onde o usuário reside ou se encontra naquele momento, que traz conteúdos como disponibilidade sobre vacinas, notícias, eventos, publicações etc. Caso o usuário relate sintomas graves, o paciente é exortado a receber atendimento médico de urgência e ligar para 911. A ferramenta STOP COVID solicita que o paciente informe a idade selecione quais comorbidades possui (doença pulmonar crônica, insuficiência cardíaca, neoplasia, doença renal crônica, insuficiência renal, diabetes e obesidade) ou ainda reside em instituição de longa permanência ou se faz uso de drogas que reduzem a imunidade. Em seguida o usuário é perguntado sobre a presença de febre (e também qual o grau), tosse e falta de ar bem como sobre a velocidade e progressão dos sintomas e se há presença de respiração ofegante ou tosse com muco. Na continuidade o paciente é indagado sobre a presença de fadiga, dor muscular, calafrio, cefaleia, diarreia, náusea, faringite, alteração de paladar ou olfato. Além disso, solicita-se que o usuário informe os países visitados nos últimos 14 dias. Ao final do preenchimento, o paciente recebe

informação quanto ao seu risco de infecção (alto, baixo) e orientações quanto a monitoramento dos sintomas, higiene, minimização de contato e recomendações de uso da telemedicina.

Tabela 1 – Características de ferramentas utilizadas para *screening* de COVID-19

Ferramenta	Idioma	Disponibilidade	Estimativa de probabilidade	Vantagens	Desvantagens
Corona Check	inglês / alemão	IoS	não	não solicita dados pessoais	mistura de idiomas no aplicativo (inglês e alemão)
C-Screener	inglês	Android e IoS	não	disponível para Android	a) solicita dados pessoais b) usuário precisar informar sim ou não para cada sintoma avaliado
Apple Covid-19	português	IoS	não	a) lista maior de sintomas considerados como ameaçadores a vida e encaminhamento para emergência b) dicas sobre como manter a saúde mental	preenchimento longo
BeebeHealth Care	Inglês	Android e IoS	não	a) lista maior de sintomas considerados como ameaçadores a vida e encaminhamento para emergência b) orientações sobre locais de realização do teste para COVID-19	a) lista grande comorbidades a serem consideradas b) lista pequena de sintomas da doença
CDC	inglês	IoS	não	a) lista maior de sintomas considerados como	preenchimento longo

					ameaçadores a vida e encaminhamento para emergência
					b) orientações sobre vacinas nas proximidades do usuário
STOP COVID	inglês	IoS	não	usuário pode fazer um diário sobre a evolução de sua doença	não traz encaminhamento para emergência se sintomas graves

2 OBJETIVOS

2.1 Geral

- a) Desenvolver um algoritmo de apoio à triagem utilizando um modelo de redes neurais artificiais (DL) para avaliar a probabilidade do diagnóstico de COVID-19 a partir dos dados clínicos.

2.2 Específicos

- a) Coletar e processar informações disponíveis sobre os sintomas da doença a partir de bancos de dados públicos e gratuitos bem como de artigos científicos, o quais serão utilizados como insumo para o modelo de DL.
- b) Desenvolver uma plataforma *online* em Javascript na qual os pacientes possam informar seus sintomas e serem informados sobre a probabilidade de o diagnóstico ser COVID-19 pelo modelo de DL.
- c) Desenvolver um módulo dentro da plataforma para fornecer informações sobre a doença bem como orientações em função da probabilidade obtida.

3 MATERIAIS E MÉTODOS

O projeto foi desenvolvido em três fases. A primeira trata da coleta e processamento dos dados que serão utilizados para treinamento, validação e teste das redes neurais artificiais bem como da construção do modelo per si. Na segunda fase, foi construída uma plataforma (“online”) para o qual o paciente insere suas informações clínicas e recebe um *feedback* a respeito da probabilidade de diagnóstico de COVID-19. Na última fase, foi desenvolvido um módulo na plataforma contendo informações para os pacientes a respeito da doença e orientações em função da probabilidade calculada pelo algoritmo.

Fase 1: Coleta e processamento de dados e desenvolvimento do modelo de ML

No âmbito desse trabalho, foi desenvolvido um modelo preditor de risco bayesiano associado a uma rede neural. Ele utiliza como atributos características clínicas do paciente e, opcionalmente, com características de imagem de sua radiografia de tórax. As informações utilizadas para treinamento dos modelos de ML foram obtidas de artigos científicos, como por exemplo, a seleção dos sintomas mais comuns e a presença de outras características clínicas importantes bem como as prevalências de cada uma encontradas na literatura (Bi et al., 2020; Yan et al., 2020; Phua et al, 2020; Huang et al., 2019; Chan et al. 2019; Yu et al. 2019). Além disso, bancos de dados públicos e gratuitos também foram consultados, como por exemplo, o tradicional site da Kaggle e o repositório COVID *DataSharing* (Mello et al, 2020). Em relação à radiografia, uma rede neural do tipo convolucional foi treinada para aprender características pulmonares anormais a partir de radiografias disponibilizadas em banco de dados público (Irvin et al., 2020). O envio da radiografia de tórax, no entanto, é opcional visto que na grande maioria das vezes o usuário não tem uma radiografia de tórax à disposição para submetê-la a avaliação do algoritmo. Utilizou-se um processo de aprendizado supervisionado, sendo 80% utilizado no treinamento do modelo, 10% na validação e o restante para teste. A avaliação do desempenho do modelo foi realizada por meio de um conjunto de métricas tradicionais que incluem: sensibilidade, especificidade, acurácia e da área abaixo da curva característica de operação do Receptor (AUC). Os modelos foram implementados nas linguagens Python e Julia, escolhidas por serem gratuitas, *open source* e de fácil utilização.

Fase 2: Desenvolvimento da plataforma online

Após o treinamento do modelo na Fase 1, este foi utilizado no desenvolvimento de uma plataforma online. Essa plataforma pode ser consultada tanto em português quanto em inglês. Ela permite que o paciente informe os sintomas apresentados nos últimos 15 dias bem como seus dados clínicos e receba como retorno qual a probabilidade de ter o diagnóstico de COVID-19. Valores de probabilidade inferiores a 50% foram considerados baixa probabilidade, entre 50 e 69% média e maiores ou iguais a 70% alta probabilidade. A plataforma não solicita nem armazena informações de cadastro dos usuários nem coleta informações dos aparelhos e dispositivos utilizados para acesso ao site e à aplicação. Entretanto, é solicitado que o usuário concorde com os seus termos de uso e privacidade, os quais incluem o propósito da ferramenta, suas limitações e *compliance* com a Lei Geral de Proteção de Dados (LGPD). No tocante a implementação, o *layout* foi desenvolvido tendo como base a estratégia *mobile first* (Firtmann, 2013). Além disso, o desenvolvimento seguiu princípios básicos da literatura de desenvolvimento de *software*, como o *clean code* (Martin, 2008) e SOLID. O *clean code* tem subsídio na busca de técnicas simples para tornar fácil tanto a leitura quanto a escrita e compressão do código. SOLID é um acrônimo para 5 princípios importantes de *design* de código, criados por Robert C. Martin: **S** (Single Responsibility – Princípio da Responsabilidade Única), **O** (Open/Closed – Princípio aberto/fechado), **L** (Liskov Substitution- Princípio da Substituição de Liskov), **I** (Interface Segregation – Princípio da Segregação de Interfaces) e **D** (Dependency Inversion – Princípio da Inversão de Dependências (Martin, 2018). O uso desses princípios permite maior facilidade de manutenção, adaptação e alterações ao escopo do projeto. Além disso, foram seguidos princípios básicos de segurança visando aumentar a confiança e proteção dos usuários, como por exemplo, o uso de *Secure Socket Layer* (SSL), que é um dos protocolos mais populares usados na internet com o intuito de facilitar uma comunicação segura através de processos de autenticação e criptografia (Alnatheer, 2014).

Fase 3: Desenvolvimento de um módulo com informações e orientações aos pacientes

Nessa fase, seguindo os mesmos princípios da fase anterior, foi construído um módulo visando fornecer informações e orientações aos pacientes sobre a COVID-19. O conteúdo do material tem como base informações de artigos científicos, do Ministério da Saúde e da

Organização Mundial de Saúde.

Registro

Após a conclusão da última etapa final do dessa etapa, foi gerado um *hash* do código fonte da versão final do produto. Além disso, o produto foi registrado no Instituto Nacional de Propriedade Intelectual sob o número BR512022002678-3 e cujos autores são Leandro dos Guimarães Marques Alvim, Filipe Braidão do Carmo, Bruno José Dembogurski, Rodrigo de Souza Tavares, Erito Marques de Souza Filho e Alexandra Maria Monteiro Grisolia.

4 RESULTADOS

Conforme se visualiza na Figura 1, ao entrar no *link* da ferramenta o usuário acessa uma tela inicial, que possui a opção de clicar para saber mais informações sobre o sistema (botão “Saiba Mais”) – que traz informações sobre o que é projeto, sua motivações e especificidades sobre a parte de programação ou iniciar a inserção de suas informações clínicas (botão “Iniciar”).

Figura 1 – Tela de apresentação do projeto



Fonte: <https://tools.atislabs.com.br/covid>

Ao clicar em “Iniciar”, conforme se vislumbra na Figura 2, o paciente tem acesso a uma tela, na qual deve selecionar os sintomas que ele apresenta nos últimos 15 dias. Além disso, na tela subsequente, como se verifica na Figura 3, ele deve informar se possui mais de 60 anos, se tem alguma comorbidade, qual seu gênero e também deve aceitar os termos e condições e a política de privacidade. Caso clique em “Anterior”, o usuário é remetido para a tela anterior aquela que está navegando.

Figura 2 – Tela de identificação e marcação dos sintomas

XRayCovid-19 **Sintomas** Radiografia Resumo

Informe os sintomas que apareceram dentro dos últimos 15 dias

escolha os sintomas abaixo

Selecione os sintomas do(a) paciente

Artralgia (Dor nas articulações)
 Vômito
 Hiposmia (Diminuição do olfato)
 Congestão nasal (Nariz entupido)
 Tosse

Escarro
 Fadiga (Cansaço)
 Hemoptise (Expectoração de sangue)
 Dispneia (Falta de ar)
 Náusea
 Diarreia

Mialgia (Dor muscular)
 Febre
 Anosmia (Perda do olfato)
 Dor abdominal
 Tonteira

Anorexia (Falta ou perda do apetite)
 Secreção ocular
 Dor de ouvido ou diminuição da audição
 Eructação (Arroto)

Calafrios
 Cefaleia (Dor de cabeça)
 Dor de garganta
 Ageusia (Perda do paladar)

Disgeusia (Alteração do paladar)

Anterior Próximo

Fonte: <https://tools.atislabs.com.br/covid>

Figura 3 – Tela de solicitação de outras informações e do aceite das políticas de privacidade

XRayCovid-19 Sintomas Radiografia **Resumo**

Informações Gerais do(a) Paciente

Possui mais de 60 anos?
 Possui alguma comorbidade?
 Gênero

Sim
 Sim
 Masculino

Não
 Não

Li e concordo com os Termos e Condições e a Política de Privacidade

Anterior Analisar

Fonte: <https://tools.atislabs.com.br/covid>

Ao clicar em “Analisar”, o algoritmo de ML desenvolvido no âmbito desse projeto retorna ao usuário qual a probabilidade de o paciente naquelas condições possuir a doença. Na Figura 4 tem-se um exemplo de um usuário com menos de 60 anos, sem comorbidades e que apresentou os sintomas de artralgia, hiposmia, tosse, febre e calafrios. A ferramenta indicou uma probabilidade de 80% de o um usuário com essas características clínicas possuir a doença.

Figura 4 – Tela de saída com a estimativa de probabilidade



Fonte: <https://tools.atislabs.com.br/covid>

5 DISCUSSÃO

No presente projeto foi desenvolvido um algoritmo de suporte a tomada de decisão capaz de avaliar a probabilidade de diagnóstico de COVID-19 em pacientes com suspeita da doença. Do melhor de nosso conhecimento, esse é a primeira aplicação em língua portuguesa que apresenta uma avaliação probabilística da suspeita da doença utilizando modelos de ML exclusivamente a partir dos sintomas e de características clínicas dos usuários. Essa probabilidade foi considerada elevada se maior ou igual a 70%, média se superior a 50% e inferior a 70% e baixa caso contrário. Todas as aplicações avaliadas no âmbito desse trabalho não fizeram uso de modelo de ML e não fizeram uma abordagem probabilística. A aplicação de ferramentas para estimar a probabilidade de uma doença tem sido de grande valia em diversos contextos clínicos distintos. Um estudo recém publicado por Liu et al. (2020), por exemplo, utilizou com sucesso redes neurais para estimar a probabilidade de Acidente Vascular Encefálico (AVE) de natureza isquêmica a partir de dados clínicos. Nesse contexto, esse tipo de ferramenta probabilística não havia sido, até a presente aplicação, utilizada no âmbito da infecção por Coronavírus Covid-19 para triagem automatizada a partir de sintomas e dados clínicos. Outro ponto importante foi a dispensa de utilização de informações obtidas por exames complementares. Ademais, o sistema web desenvolvido possui como vantagem, conforme pontuado por Handl et al. (2020), a possibilidade de livre utilização (sem custo) pelas pessoas, desde que possuam acesso à Internet. Diferentemente de outras ferramentas, como se verifica no C-Screener, a ferramenta desenvolvida neste projeto também tem como vantagem a não necessidade de preenchimento, por parte do usuário, de informações pessoais como nome, telefone e email - o acrescenta uma segurança adicional com respeito a anonimização do usuário e segurança de dados sensíveis. O usuário do sistema *web* tem a opção de saber mais a respeito do trabalho e em seguida selecionar quais os sintomas possui. Os sintomas são apresentados tanto pela nomenclatura formal quanto pelo nome popular, diferentemente de outras plataformas disponíveis. Além disso, o usuário deve concordar com os termos e condições e a política de privacidade - prática importante, conforme destacado por Mbunge (2020) e que se alinha as necessidades de transparência e *compliance* ressaltadas por Lucivero et al. (2020). Essas características se alinham as diretrizes preconizadas na LGPD, já em vigor, com relação a proteção de dados dos usuários do sistema.

Ademais, algumas das ferramentas avaliadas trazem informações sobre sintomas de

maior gravidade (*Apple COVID-19*, *Beebe HealthCare* e CDC), indicando a necessidade de comparecimento a emergência, o que não é contemplado ainda na versão desenvolvida no âmbito desse trabalho. Dessa forma, tem grande utilidade para os usuários a construção de um módulo com orientações que inclua esse tipo de informação, desde que respeitando os limites da ferramenta desenvolvida. Com relação aos idiomas das aplicações, apenas o *Apple COVID-19* encontra-se disponível em português. A ferramenta desenvolvida aqui neste projeto tem a vantagem de ser bilíngue. Por outro lado, a escolha de todos os sintomas a partir de uma tela única permite maior usabilidade no tempo de preenchimento quando se compara a um conjunto de várias telas sequenciais, como por exemplo, na ferramenta desenvolvida pelo CDC e no *Beebe HealthCare*.

Alem disso, é importante o uso de outras ferramentas como o uso do *Google Analytics* (GA) para avaliação de estatísticas de acesso e outras características dos usuários – informações que podem ser utilizadas por pesquisadores e gestores de saúde pública. Uma experiência bem sucedida com o uso do GA (e de outras ferramentas) foi obtida por Gesualdo et al. (2020) com o intuito de disponibilizar acesso a informações de qualidade sobre segurança de vacinas. A aplicação desenvolvida no âmbito desse trabalho é capaz de fornecer esse tipo de dado anonimizado.

CONCLUSÃO

Este projeto desenvolveu uma ferramenta bilingue para triagem de pacientes com sinais e sintomas suspeitos para a doença provocada pelo Coronavírus COVID-19. Ela tem por diferencial o cálculo automatizado da probabilidade de risco da infecção propriamente dito utilizando um modelo de ML. Trata-se de um projeto inovador e pioneiro que utiliza como insumo dados rotineiros e de fácil obtenção. Além disso, traz no seu bojo a relevância dos modelos de Inteligência Artificial no âmbito da prática médica tanto no suporte a tomada de decisão quanto em relação ao empoderamento dos pacientes, em particular no tocante a ampliação de sua literacia em saúde.

REFERÊNCIAS

- 1 – Alnatheer MA. Secure Socket Layer (SSL) Impact on Web Server Performance. *Journal of Advances in Computer Networks*, Vol. 2, No. 3, September 2014
- 2 – Bi Q, Wu Y, Mei S, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study [published online ahead of print, 2020 Apr 27]. *Lancet Infect Dis.* 2020;S1473-3099(20)30287-5. doi:10.1016/S1473-3099(20)30287-5
- 3 – Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm>. Acesso em: 08 ago. 2020.
- 4 – Breiman L. Random forests. *Mach Learn.* 2001;45(1):5-32.
- 5 – Broomhead DH, Lowe D. Multivariable Functional Interpolation and Adaptive Networks (1988). *Complex Systems.* 2: 321–355.
- 6 – Chan JF, Yuan S, Kok KH, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet.* 2020;395(10223):514- 523. doi:10.1016/S0140-6736(20)30154-9
- 7 – Coleman JJ, Manavi K, Marson EJ, Botkai AH, Sapey E. COVID-19: to be or not to be; that is the diagnostic question. *Postgrad Med J.* 2020;96(1137):392-398. doi:10.1136/postgradmedj-2020-137979
- 8 – Cortes C, Vapnik V. Support-vector networks. *Mach Learn.* 1995;20:273-97.
- 9 – Flanagan, David; Ferguson, Paula (2002). *JavaScript: The Definitive Guide* 4th ed. [S.l.]: O'Reilly & Associates. ISBN 0-596-00048-0
- 10 – Firtman M. *Programming the Mobile Web.* (2013). pp. 512. O'Reilly Media. ISBN 978-

0-596-80778-8.

11 – Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat.* 2001;29(5):1189-232.

12 – Gesualdo F, Marino F, Mantero J, Spadoni A, Sambucini L, Quaglia G, Rizzo C, Sahinovic I, Zuber PLF, Tozzi AE; VSN-WAP study group. The use of web analytics combined with other data streams for tailoring online vaccine safety information at global level: The Vaccine Safety Net's web analytics project. *Vaccine.* 2020 Sep22;38(41):6418-6426. doi: 10.1016/j.vaccine.2020.07.070. Epub 2020 Aug9. PMID: 32788137.

13 – Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med.* 2020;382(18):1708-1720.

14 – Hair MG, Nobre FF, Brasil P. Characterization of clinical patterns of dengue patients using an unsupervised machine learning approach. *BMC Infect Dis.* 2019 Jul 22;19(1):649. doi: 10.1186/s12879-019-4282-y. PMID: 31331271; PMCID: PMC6647280.

15 – Hastie T, Tibshirani R, Friedman JH. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* 2nd ed. New York, NY: Springer; 2009

16 – Hendl T, Chung R, Wild V. Pandemic Surveillance and Racialized Subpopulations: Mitigating Vulnerabilities in COVID-19 Apps. *J Bioeth Inq.* 2020 Aug 25:1–6.doi: 10.1007/s11673-020-10034-7. Epub ahead of print. PMID: 32840858; PMCID: PMC7445800.

17 – Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China [published correction appears in *Lancet.* 2020 Jan 30;:]. *Lancet.* 2020;395(10223):497- 506. doi:10.1016/S0140-6736(20)30183-5

18 – Irvin J, Rajpurkar P, Ko M et al. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. arXiv preprint arXiv:1901.07031 (2019).

19 – James G, Witten D, Hastie T, Tibshirani R. An Introduction to Statistical Learning with Applications in R. Springer Texts in Statistics, Springer Science+Business Media, New York 2013.

20 – LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-444.
doi:10.1038/nature14539

21 – Liu Y, Yin B, Cong Y. The Probability of Ischaemic Stroke Prediction with a Multi-Neural-Network Model. Sensors (Basel). 2020 Sep 3;20(17):4995. doi: 10.3390/s20174995. PMID: 32899242; PMCID: PMC7506623.

22 – Lucivero F, Hallowell N, Johnson S, Prainsack B, Samuel G, Sharon T. COVID-19 and Contact Tracing Apps: Ethical Challenges for a Social Experiment on a Global Scale. J Bioeth Inq. 2020 Aug 25:1–5. doi: 10.1007/s11673-020-10016-9. Epub ahead of print. PMID: 32840842; PMCID: PMC7445718.

23 – Martin RC. Clean Code A Handbook of Agile Software Craftsmanship. (2008). pp. 431. Prentice Hall. ISBN 0-13-235088-2

24 – Martin RC. Agile Software Development: Principles, Patterns, and Practices. (2013). pp. 524. Pearson. ISBN 978-1-292-02594-0

25 – Martin R (2018). Clean Architecture: A Craftsman's Guide to Software Structure and Design. Prentice Hall. ISBN 9780134494166.

26 – Mbunge E. Integrating emerging technologies into COVID-19 contact tracing: Opportunities, challenges and pitfalls. Diabetes MetabSyndr. 2020 Aug 26;14(6):1631-1636. doi: 10.1016/j.dsx.2020.08.029. Epub ahead of print. PMID: 32892060.

27 – McCarthy J, Minsky ML, N. Rochester, C.E. Shannon. August 31, 1955. A proposal for the Dartmouth Summer Research Project on artificial intelligence. Disponível em: <<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>>. Acesso em: 29 out.

2020.

28 – McCorduck P. (2004). *Machines Who Think: a Personal Inquiry into the History and Prospects of Artificial Intelligence* (2nd ed.), Natick, MA: A. K. Peters, Ltd. ISBN 1-56881205-1.

29 – McCulloch W, Pitts W. A Logical Calculus of Ideas Immanent in Nervous Activity. 1943. *Bulletin of Mathematical Biophysics* 5:115–133.

30 – Mello, Luiz E., Suman, Andrea, Medeiros, Claudia Bauzer, Prado, Claudio Almeida, Rizzatti, Edgar Gil, Nunes, Fatima L. S., ... Carvalho, Rodrigo. (2020, July 30). Opening Brazilian COVID-19 patient data to support world research on pandemics. Zenodo. <http://doi.org/10.5281/zenodo.3966427>

31 – Ministério da Saúde. Diretrizes para diagnósticos e tratamento da COVID-19. 2020. Disponível em: <<https://portalarquivos.saude.gov.br/images/pdf/2020/April/18/Sumario-Covid19.pdf>>. Acesso em: 29 out. 2020.

32 – Minsky M. L. and Papert S. A. 1969. *Perceptrons*. Cambridge, MA: MIT Press.

33 – Ngiam KY, Khor IW. Big data and machine learning algorithms for health-care delivery. *Lancet Oncol.* 2019 May;20(5):e262-e273. doi: 10.1016/S1470-2045(19)30149-4. Erratum in: *Lancet Oncol.* 2019 Jun;20(6):293. PMID: 31044724.

34 – Nicola M, Alsafi Z, Sohrabi C, et al. The socio-economic implications of the coronavirus pandemic (COVID-19): A review. *Int J Surg.* 2020;78:185-193. doi:10.1016/j.ijssu.2020.04.018

35 – Phua J, Weng L, Ling L, et al. Intensive care management of coronavirus disease 2019 (COVID-19): challenges and recommendations [published correction appears in *Lancet Respir Med.* 2020 May;8(5):e42]. *Lancet Respir Med.* 2020;8(5):506- 517. doi:10.1016/S2213-2600(20)30161-2

36 – Richardson S, Hirsch JS, Narasimhan M, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area [published online ahead of print, 2020 Apr 22]. *JAMA*. 2020;e206775.

doi:10.1001/jama.2020.6775

37 – Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. (1958) Cornell Aeronautical Laboratory, *Psychological Review*, v65, No. 6, pp. 386–408.

38 – Rumelhart DE; Hinton, Geoffrey E.; Williams, Ronald J. (8 October 1986). Learning representations by back-propagating errors. *Nature*. 323 (6088): 533–536.

39 – Souza Filho EM, Fernandes FA, Soares CLA, et al. Artificial Intelligence in Cardiology: Concepts, Tools and Challenges - "The Horse is the One Who Runs, You Must Be the Jockey". *Arq Bras Cardiol*. 2020;114(4):718-725. doi:10.36660/abc.20180431

40 – Than MP, Pickering JW, Sandoval Y, Shah ASV, Tsanas A, Apple FS, Blankenberg S, Cullen L, Mueller C, Neumann JT, Twerenbold R, Westermann D, Beshiri A, Mills NL; MI3 collaborative. Machine Learning to Predict the Likelihood of Acute Myocardial Infarction. *Circulation*. 2019 Aug 16;140(11):899–909. doi: 10.1161/CIRCULATIONAHA.119.041980. Epub ahead of print. PMID: 31416346; PMCID: PMC6749969.

41 – Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*. 2020;382(8):727-733.

42 – Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China [published online ahead of print, 2020 Feb 7]. *JAMA*. 2020;323(11):1061-1069. doi:10.1001/jama.2020.1585

43 – Ye Z, Zhang Y, Wang Y, Huang Z, Song B. Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review [published online ahead of print, 2020 Mar 19]. *EurRadiol*. 2020;1-9. doi:10.1007/s00330-020-06801-0

44 – Yu X, Sun X, Cui P, et al. Epidemiological and Clinical Characteristics of 333 Confirmed Cases with Coronavirus Disease 2019 in Shanghai, China [published online ahead of print, 2020 Apr 29]. *Transbound Emerg Dis.* 2020;10.1111/tbed.13604.
doi:10.1111/tbed.13604

45 – Yan Y, Yang Y, Wang F, et al. Clinical characteristics and outcomes of patients with severe covid-19 with diabetes. *BMJ Open Diabetes Res Care.* 2020;8(1):e001343.
doi:10.1136/bmjdr-2020-001343

Sítios consultados

- 1 – Kaggle. Disponível em: <https://www.kaggle.com>
- 2 – Python. Disponível em: <https://www.python.org/>
- 3 – Julia. Disponível em: <https://julialang.org/>
- 4 – Google Analytics. Disponível em: <https://analytics.google.com/analytics/web/>
- 5 – Ministério da Saúde: disponível em: <https://coronavirus.saude.gov.br/sobre-a-doenca>
- 6 – CoronaCheck: disponível em <https://coronacheck.nl/nl/>
- 7 – C-Screener: disponível em <https://apps.apple.com/br/app/c-screener/id1535157998>
- 8 – Apple Covid-19: disponível em <https://www.apple.com/newsroom/2020/03/apple-releases-new-covid-19-app-and-website-based-on-CDC-guidance/>
- 9 – BeebeHealthCare: disponível em <https://www.beebehealthcare.org/>
- 10 – CDC: disponível em <https://www.cdc.gov/>
- 11 – STOP COVID: disponível em <https://www.nidirect.gov.uk/articles/coronavirus-covid-19-stopcovid-ni-proximity-app>