



Estratégias Preditivas na Detecção do Agravamento do Quadro Clínico de Pacientes com COVID-19: Uma Revisão Sistemática da Literatura

Wallace Duarte de Holanda¹, Lenardo Chaves e Silva²

¹Programa de Pós-Graduação em Ciências da Computação - PPgCC
UERN/UFERSA, Mossoró, RN, Brasil

²Departamento de Engenharias e Tecnologia - DETEC
UFERSA, Pau dos Ferros, RN, Brasil

wallace.holanda@alunos.ufersa.edu.br, lenardo@ufersa.edu.br

Abstract. *COVID-19 is a disease caused by the new coronavirus, called SARS-COV-2. This disease has a varied clinical aspect, which makes it difficult to detect and risk getting worse. In this context, numerous competent studies seeking the development of models and predictive strategies within the scope of COVID-19. Thus, a Systematic Literature Review was carried out, seeking to understand the current research landscape in relation to a predictive analysis of the diagnosis and the worsening capacity of patients with a COVID-19. The collected results show that predictive models based on supervised learning techniques are being developed through programming languages and analytical softwares.*

Resumo. *A COVID-19 é a doença causada pelo novo coronavírus, denominado SARS-COV-2. Esta doença possui aspecto clínico variado, o que dificulta o sua detecção e risco de agravamento. Neste contexto, numerosos estudos vêm buscando o desenvolvimento de modelos e estratégias preditivas no âmbito da COVID-19. Com isso, foi conduzida uma Revisão Sistemática da Literatura, buscando compreender o panorama atual de pesquisas em relação a análise preditiva do diagnóstico e capacidade de agravamento de pacientes com a COVID-19. Os resultados coletados apontam que modelos preditivos baseados em técnicas de aprendizagem supervisionado estão sendo desenvolvidos, por meio de linguagens de programação e softwares analíticos.*

1. Introdução

A COVID-19 é a doença causada pelo novo coronavírus, denominado SARS-CoV-2. Devido seu aspecto clínico variante, uma parcela dos infectados pode não manifestar os sintomas, enquanto outros podem desenvolver os sintomas em larga escala, impactando em casos graves. Estima-se que cerca de 80% dos pacientes com COVID-19 podem ser assintomáticos ou oligossintomáticos (poucos sintomas), e aproximadamente 20% dos casos detectados requer atendimento hospitalar por apresentarem dificuldade respiratória, dos quais aproximadamente 5% podem necessitar de suporte ventilatório [Brasil 2020].

Os sintomas da COVID-19 podem variar de um resfriado a uma Síndrome Gripal (SG). Dessa maneira, os pacientes podem vir a manifestar desde sintomas leves (febre,

tosse seca e cansaço) e incomuns (dor de garganta, diarreia, conjuntivite, dor de cabeça e perda de sabor ou cheiro) até graves (falta de ar, dor ou pressão no peito e perda de fala ou movimento) [Brasil 2020, WHO 2020].

Por conta desta característica variante, os profissionais da saúde apresentam dificuldade em compreender o comportamento da doença, tornando-se complicado a detecção e a possibilidade de agravamento. Mediante esta problemática, técnicas preditivas vêm sendo adotadas, de modo a auxiliar no processo de identificação de padrões de sintomas e características dos pacientes infectados [Shinde et al. 2020].

Segundo [Jamshidi et al. 2020] essas técnicas podem auxiliar na descoberta da relações entre os diversos parâmetros que envolvem a doença, com intuito de fornecer novas informações a partir de dados já conhecidos. [Shinde et al. 2020] enfatiza que dentre principais técnicas utilizadas, destacam-se as baseadas em modelos matemáticos, ciência de dados, aprendizagem de máquinas e inteligência artificial.

Diante desse cenário, este trabalho tem como objetivo de identificar na literatura estratégias preditivas sobre o agravamento de pacientes com COVID-19. Dessa maneira, foi definida a seguinte questão de pesquisa: Qual o panorama atual de pesquisas em relação a análise preditiva referente ao agravamento do quadro clínico de pacientes com a COVID-19?

Este artigo encontra-se organizado da seguinte forma. A Seção 2 discute alguns trabalhos relacionados com a temática deste estudo. A Seção 3 apresenta o protocolo utilizado. Na Seção 4 são apresentados os resultados atingidos. Na Seção 5 são discutidos os resultados alcançados. E, na Seção 6 são expostas as considerações finais.

2. Trabalhos Relacionados

Alguns estudos [Shamsoddin 2020, Katzenschlager et al. 2020] vêm utilizando a Revisão Sistemática de Literatura (RSL) com objetivo de obter informações sobre modelos e técnicas desenvolvidas, de modo a identificar os principais sintomas e características atreladas ao agravamento COVID-19.

Em [Shamsoddin 2020] foi conduzida uma RSL visando avaliar e examinar modelos preditivos direcionados no diagnóstico e risco de infecção do COVID-19. Para a pesquisa, foram selecionaram trabalhos entre 2019 e Maio de 2020, sendo obtido ao final, 107 artigos selecionados e 145 modelos de predição. Como resultado, foi observado que a idade, temperatura corporal, contagem de linfócitos e características de imagem pulmonares são os principais fatores em comum nos modelos preditivos. Entretanto, mediante a inconsistência dos sintomas, foi identificado que mesmo modelos com acurácia elevada ainda possuem um alto risco de viés.

Em [Katzenschlager et al. 2020] foi realizada uma RSL com intuito de analisar a semelhança entre os principais fatores de risco identificados em pacientes hospitalizados com a COVID-19. Os autores selecionaram artigos entre Dezembro de 2019 e Maio de 2020. Ao final da análise, 88 artigos foram escolhidos para serem averiguados. Os resultados mostraram que a idade avançada, presença de alguma doença cerebrovascular e o PCR são os principais indicadores na admissão de UTIs e óbito dos pacientes.

Em uma perspectiva diferente, este estudo busca identificar e avaliar trabalhos que apresentem modelos preditivos relacionados a possibilidade de agravamento do quadro

clínico de pacientes com COVID-19. Com isso, além de compreender o panorama atual de pesquisas nesta área, será possível identificar os principais modelos, técnicas e tecnologias utilizadas na construção dos modelos e origem dos dados analisados.

3. Método Utilizado

Segundo [Dresch et al. 2015] uma RSL se caracteriza como um estudo secundário, que visa mapear, encontrar, avaliar e agregar as contribuições de estudos primários em relação a uma questão de pesquisa específica. Com isso, por ser um estudo sistemático, a RSL deve seguir um protocolo planejado e fundamentado, de modo a garantir que o processo possa ser replicado e livre de viés. Para a realização desta RSL, foi utilizado o protocolo definido em [Dresch et al. 2015].

3.1. Questões de Pesquisa

A questão central de pesquisa desta RSL consiste em compreender o panorama atual de trabalhos científicos que apresentem o desenvolvimento de modelos e técnicas preditivas voltadas ao agravamento do quadro clínico de pacientes com COVID-19. Para responder esta questão central, foram elaboradas 3 questões de pesquisa específicas (QE):

- QE1: Quais as principais abordagens (modelos, algoritmos e técnicas) utilizadas para análise de dados?
- QE2: Quais recursos tecnológicos (linguagens, ferramentas e bibliotecas) foram utilizadas na análise de dados?
- QE3: Qual a origem das base de dados empregadas na análise de dados?
- QE4: Quais variáveis foram considerados mais relevantes na análise de dados?
- QE5: Quais as limitações identificadas no desenvolvimento do trabalho?

3.2. Processo de Busca

A procura por trabalhos primários foi conduzida por meio de uma busca manual em duas fontes de pesquisa na área de tecnologia: *ACM Digital Library*, *IEEE Xplore*, *MedLine*, *MDPI*, *Scopus* e *Web of Science*. Essas fontes foram escolhidas baseando-se na: i) disponibilidade em acessar o conteúdo dos artigos, ii) importância e consolidação das fontes, iii) presença de eventos e conferências voltadas para o objetivo desta pesquisa.

Para identificar as principais expressões que remetesse aos trabalhos buscados nesta pesquisa, foram definidos termos ligados a área médica, tecnológica e gerais. Os termos foram escritos em inglês, tendo em vista o conteúdo dos artigos disponíveis nas fontes de busca. A Tabela 1 apresenta dos termos selecionados.

Tabela 1. Termos de busca.

Área	Termos
Médica	COVID; COVID-19; Sars-Cov-2; Coronavirus.
Tecnológica	Machine Learning; Data Mining; Artificial Intelligence; Data Science; Deep Learning.
Geral	Model; Prediction; Diagnosis.

Para realizar o processo de busca nos anais das fontes de busca, foram selecionados um conjunto de *strings* de busca baseando-se nos termos definidos na Tabela 1. Com

isso, a seguinte *string* de busca foi elaborada: (*COVID OR COVID-19 OR Sars-Cov-2 OR Coronavirus*) *AND* (*Machine Learning OR Data Mining OR Artificial Intelligence OR Data Science OR Deep Learning*) *AND/OR* (*Model, Predict OR Diagnosis*).

3.3. Seleção e Avaliação

O processo de escolha e análise dos artigos foi dividido em quatro estágios. A pré-seleção (estágio 0) dos artigos caracterizou-se na verificação de trabalhos nas fontes de busca. Neste momento, foram examinados o título e as palavras-chave, considerando a *string* de busca definida na seção 3.2 e os critérios de inclusão (CI2 e CI3) apresentados na Tabela 2.

No estágio 1, foram observados o título, resumo e as palavras-chave de acordo com os critérios de exclusão (CE1, CE2, CE3) e inclusão (CI1, CI2, CI3). Os trabalhos incluídos foram adicionados à Lista de Incluídos (LI1), enquanto os excluídos foram organizados em uma Lista de Excluídos (LE1).

No estágio 2, foi realizada a leitura do título, resumo, palavras-chave, introdução, metodologia adotada e conclusão dos trabalhos da LI1. Para este estágio, foram utilizados os critérios de exclusão (CE1, CE2, CE3 e CE4) e inclusão (CI1, CI2 e CI3). Com isso, os estudos excluídos foram inseridos na Lista de Excluídos (LE2), enquanto os trabalhos incluídos foram selecionadas em uma Lista de Incluídos (LI2).

No estágio 3, foram analisado os trabalhos incluídos na LI2. Esses trabalhos foram analisados por completo, de acordo com todos os critérios de inclusão e exclusão. Ao fim, os trabalhos que satisfizeram todos os critérios foram direcionados a responder as questões de pesquisa específicas da pesquisa.

Tabela 2. Critérios de Inclusão e Exclusão.

Critérios de Inclusão	Critérios de Exclusão
CI1: Estudos com foco no objetivo da pesquisa.	CE1: Estudos repetidos em mais de uma fonte de busca.
CI2: Publicados entre os anos de 2019 e 2020 (Novembro).	CE2: Estudos sem acesso disponível.
CI3: Estudos publicados no idioma Inglês.	CE3: Estudos que não utilizem dados estruturados na análise.
CI4: Estudos que abordem pelo menos um modelo estatístico/matemático e/ou de predição.	CE4: Estudos que não respondem a nenhuma das questões de pesquisa.

3.4. Avaliação de Qualidade

A avaliação da qualidade dos trabalhos selecionados ao final do estágio 3 foi realizada a partir de uma pontuação final atribuída a cada estudo, com base nos critérios de qualidade. Os critérios foram definidos a partir de características relevantes, tais como: relevância dos resultados às questões de pesquisa, local de publicação, acessibilidade dos dados e clareza da metodologia. Nesta RSL foram utilizadas quatro critérios de qualidade:

- CQ1: Os resultados apresentados são completos e respondem às questões de pesquisa? (1 Ponto para cada pergunta respondida);

- CQ2: O trabalho foi publicado em algum meio relevante? (0 - Não; 1 - Conferência; 2 - Capítulo de livro; 3 - *Journal* com Fator de Impacto até 3,12; 4 - *Journal* com Fator de Impacto maior que 3,12);
- CQ3: A Base de Dados utilizada é acessível? (0 - Não; 1 - Sim);
- CQ4: A metodologia é clara de modo a possibilitar a replicação do trabalho? (0 - Não; 1 - Parcialmente; 2 - Totalmente).

4. Resultados

Nesta seção serão expostos os resultados obtidos ao término da realização dos estágios da RSL, com intuito de organizar as informações alcançadas e fornecer um parecer sobre as questões de pesquisa específicas.

4.1. Informações Gerais

A busca realizada resultou em 329 estudos primários^a. Ao final da execução do Estágio 1^b, 60 trabalhos foram selecionados. Com o término do Estágio 2^c, 12 trabalhos foram apurados. Ao fim do Estágio 3^d, 9 trabalhos foram incluídos, de modo a fornecer respostas às QEs. A Tabela 3 apresenta a quantidade de artigos obtidos por fonte de busca, ressaltando o total e a porcentagem de trabalhos incluídos.

Tabela 3. Número de trabalhos por fonte de busca.

Fonte	Estágio 0	Estágio 1	Estágio 2	Incluídos (%)
ACM	46	3	0	0 (0%)
IEEE Xplore	65	5	0	0 (0%)
MedLine	53	4	4	4 (7.5%)
MDPI	24	0	0	0 (0%)
Scopus	79	14	2	0 (0%)
Web of Science	62	34	6	5 (8.0%)
Total	329	60	12	9 (2.7%)

4.2. Trabalhos Selecionados

Para a obtenção das respostas às questões de pesquisa, os trabalhos selecionados [Abdulaal et al. 2020, Assaf et al. 2020, Booth et al. 2020, Cheng et al. 2020, Das et al. 2020, Kim et al. 2020, Pan et al. 2020, Parchure et al. 2020, Vaid et al. 2020] foram analisados por completo. A Tabela 4 apresenta o ID atribuído, autores, fonte de busca e título dos trabalhos analisados.

^a Artigos analisados no Estágio 0: <https://cutt.ly/Zhhkav6>

^b Artigos analisados no Estágio 1: <https://cutt.ly/CzYYdom>

^c Artigos analisados no Estágio 2: <https://cutt.ly/6zYUhc>

^d Artigos analisados no Estágio 3: <https://cutt.ly/mzYUm3i>

Tabela 4. Descrição dos Artigos Selecionados.

ID	Autores	Fonte	Título
3	[Booth et al. 2020]	Web of Science	Development of a prognostic model for mortality in COVID-19 infection using machine learning
7	[Das et al. 2020]		Predicting CoVID-19 community mortality risk using machinelearning and development of an online prognostic tool
22	[Abdulaal et al. 2020]		Prognostic Modeling of COVID-19 Using Artificial Intelligence in the United Kingdom: Model Development and Validation
25	[Assaf et al. 2020]		Utilization of machine-learning models to accurately predict the risk for critical COVID-19
46	[Cheng et al. 2020]		Using Machine Learning to Predict ICU Transfer in Hospitalized COVID-19 Patients
266	[Vaid et al. 2020]	MedLine	Machine Learning to Predict Mortality and Critical Events in a Cohort of Patients With COVID-19 in New York City: Model Development and Validation
287	[Pan et al. 2020]		Prognostic Assessment of COVID-19 in the Intensive Care Unit by Machine Learning Methods: Model Development and Validation
293	[Kim et al. 2020]		An Easy-to-Use Machine Learning Model to Predict the Prognosis of Patients With COVID-19: Retrospective Cohort Study
297	[Parchure et al. 2020]		Development and validation of a machine learning-based prediction model for near-term in-hospital mortality among patients with COVID-19

4.2.1. QE1. Quais as principais abordagens (modelos, algoritmos e técnicas) utilizadas para análise de dados e construção dos modelos?

Dentre as principais técnicas e algoritmos utilizados, destacam-se: *Logistic Regression* (ID7), *Neural Net* (ID22, ID25), *Random Forest* (ID46 e ID297), *SVM* (ID3) e *XGBoost* (ID266, ID287 e ID293). Com isso, foi observado que somente técnicas de aprendizagem direcionadas na abordagem supervisionada foram utilizadas, tendo em vista a natureza classificatória do problema. A Figura 1 apresenta as técnicas encontradas juntamente

com a quantidade de estudos que a utilizaram.

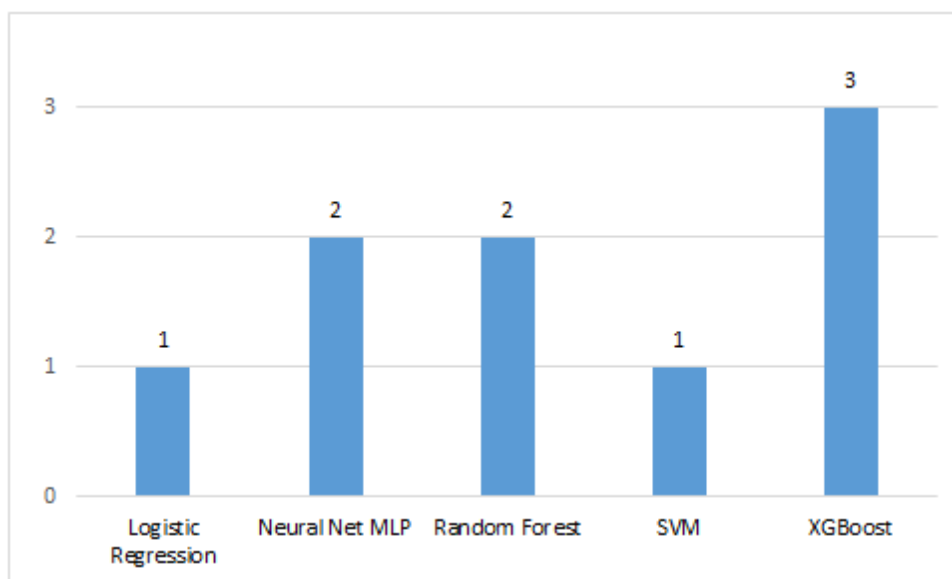


Figura 1. Técnicas utilizadas no desenvolvimento dos Modelos.

Ao analisar os trabalhos, notou-se uma forte similaridade na maneira como os estudos eram conduzidos, pois, inicialmente os autores realizavam a seleção de um conjunto de técnicas para serem testadas mediante a base de dados, e ao fim, a técnica que obtivesse melhor desempenho era escolhida para compor o modelo preditivo. Além disso, vale ressaltar que nenhum dos trabalhos selecionados fez uso de duas ou mais técnicas na construção do modelo de predição.

4.2.2. QE2. Quais recursos tecnológicos (linguagens, ferramentas e bibliotecas) foram utilizadas na análise de dados?

Baseando-se nos artigos selecionados, foram catalogadas 5 linguagens/ferramentas utilizadas na análise dos dados e criação dos modelos, a citar a linguagem Python (ID3, ID7, ID22, ID266 e ID287), linguagem R (ID7, ID46 e ID293), *software* IBM SPSS (ID25), *software* AutoML H20.ai (ID293) e *software* Apache Spark (ID46 e ID297). A Figura 2 apresenta a distribuição das linguagens e ferramentas pela quantidade de vezes empregadas.

Mediante o uso de linguagens de programação, alguns autores destacaram quais bibliotecas e pacotes foram utilizados durante o desenvolvimento, a citar (ID3, ID22, ID266 e ID287) que fizeram uso da linguagem Python em conjunto com a biblioteca de algoritmos de aprendizagem de máquina *scikit-learn*^e, (ID46 e ID293) que utilizaram a linguagem R juntamente com o pacote *PRROC*^f, voltado a visualização das métricas de precisão e revocação em formato de curva, e (ID3) que empregaram o pacote *Shiny*^g

^e<https://cutt.ly/XhhJ11B>

^f<https://cutt.ly/AhhJ3fO>

^g<https://cutt.ly/Pxw43GR>

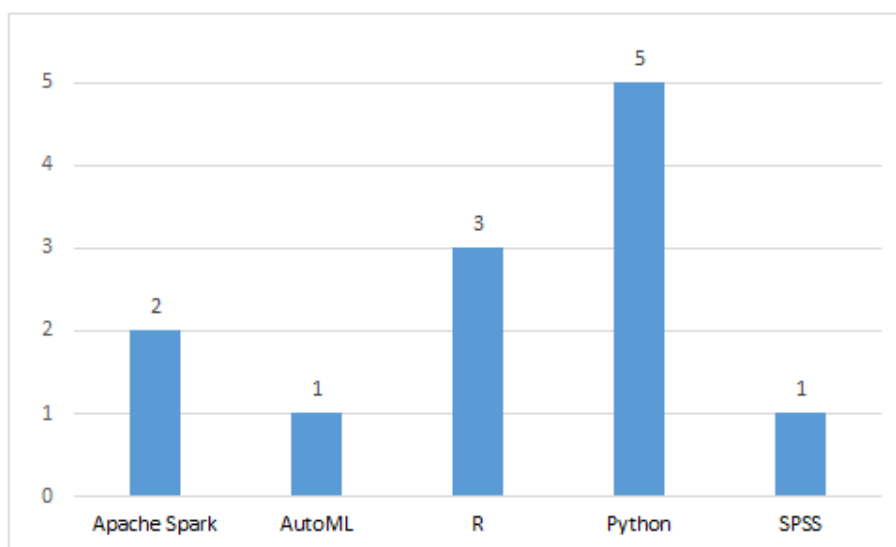


Figura 2. Ferramentas e linguagens utilizadas na criação dos Modelos.

advindo da linguagem R, possibilitando a disponibilização de modelos e aplicações no âmbito *web*.

Além disso, notou-se que alguns estudos fizeram uso de mais de uma linguagem ou tecnologia, com intuito de se beneficiar de seus pacotes e bibliotecas, a mencionar: i) (ID7)^h, englobando as linguagens Python e R; ii) (ID46), utilizando o *software* Apache Spark e a linguagem R, e iii) (ID293), aplicando o AutoML em conjunto com a linguagem R.

4.2.3. QE3. Quais as base de dados foram empregadas na análise de dados?

Para o treinamento dos modelos desenvolvidos, os autores utilizaram informações advindas de entidades médicas, destacando-se a origem de: Dados governamentais abertos (ID7)^h; ii) clínicas e centros médicos (ID3 e ID25), e iii) Hospitais (ID22, ID46, ID266, ID287, ID293 e ID297). Mediante essa definição, foi constatado que nenhum dos trabalhos selecionados utilizou bases abertas que possibilitem a visualização das informações.

Ao verificar os trabalhos que utilizaram bases de dados provenientes de hospitais e clínicas locais, foi constatado a necessidade de um maior tempo gasto na etapa de pré-processamento, tendo em vista a presença de dados faltosos e duplicados. Além disso, os autores destacaram que esta fato, se não bem tratado pode vir a influenciar os resultados.

4.2.4. QE4: Quais variáveis foram consideradas mais relevantes na análise de dados?

Mediante o desenvolvimento dos modelos preditivos, os autores conseguiram identificar os principais atributos que maior cooperaram na possibilidade do agravamento do quadro clínico dos pacientes. Com isso, mesmo com estudos utilizando um conjunto de variáveis

^hBase utilizada: <http://www.cdc.go.kr/index.es?sid=a3>

diferentes para a condução do estudo, foi possível constatar similaridades entre as características descobertas. Para ilustrar os fatores atrelados a possibilidade de agravamento, foram selecionados os 3 principais atributos de cada estudo, resultando na Tabela 5.

Tabela 5. Principais atributos relacionados ao agravamento.

Atributo	Quantidade	Atributo	Quantidade
Idade	6	Estado da Função Renal	1
Dipneia	2	Hiato Aniônico	1
Contagem de Linfócitos	2	Histórico de Fumante	1
Proteína C Reativa	2	Localidade	1
Sexo	2	Estado Mental Alterado	1
Cálcio Sérico	1	Nitrogênio Ureico	1
Contagem de Leucócitos	1	Saturação de Oxigênio	1
CRP	1	Nível de Sódio	1
Dias de Infecção	1	Tempo de Protrombina	1

Devido a variabilidade de dados utilizados no desenvolvimento dos trabalhos selecionados, a maioria dos atributos tiveram aparição em poucos estudos. Entretanto, a idade avançada, dispneia, contagem de linfócitos, proteína C reativa e o sexo do paciente foram relatados como variáveis importantes, sendo destacadas em estudos diferentes. Mediante essa panorama, é possível compreender que devido seu aspecto clínico instável, um conjunto de características distintas podem estar envolvidas na possibilidade do agravamento.

4.2.5. QE5: Quais as limitações identificadas no desenvolvimento do trabalho?

Mediante a realização desta RSL, foi possível identificar que alguns trabalhos, relataram problemas nos procedimentos envolvidos na construção dos modelos preditivos. A citar: i) a baixa quantidade de amostras (instâncias) presentes nas base de dados selecionadas na composição do modelo (ID3, ID22, ID25 e ID287); ii) escassez de atributos, mesmo em bases com muitas instâncias (ID7); iii) número desequilibrado de pacientes que vieram a óbito ou se recuperaram (ID46 e ID293), e iv) impossibilidade da realização de testes reais em hospitais e clínicas através do modelo desenvolvido (ID297). A Tabela 6 apresenta a disposição das características das bases utilizadas relacionadas a: i) quantidade de instâncias, i) número de atributos, e iii) data de coleta.

Mediante a ilustração da Tabela, foi possível compreender que os 4 trabalhos (ID3, ID22, ID25 e ID287) que destacaram a ausência de instâncias para análise dos dados, utilizaram bases com quantidades similares de instâncias, comprovando a baixa quantidade amostras de existentes. Além disso, o artigo (ID7) que um dos pontos fracos do estudo se resume na escassez de atributos, fato esse, que pode ser constatado, tendo em vista que apenas 5 variáveis foram utilizadas na construção do modelo.

5. Discussão dos Resultados

A partir dos resultados alcançados, nota-se que a busca por estratégias preditivas relacionadas a probabilidade de agravamento do quadro clínico de pacientes com COVID-19, mesmo com poucos meses do seu descobrimento, vem se denotando como uma tema

Tabela 6. Características das Bases de Dados

ID	Instâncias	Atributos	Data de Coleta
3	398	26	Não Informado
7	3524	5	20/01/2020 - 30/05/2020
22	398	22	02/02/2020 - 22/04/2020
25	162	23	08/03/2020 - 05/04/2020
46	1987	31	26/02/2020 - 18/04/2020
266	4098	45	15/03/2020 - 22/05/2020
287	123	100	02/02/2020 - 15/04/2020
293	4787	35	25/01/2020 - 03/06/2020
297	567	49	10/02/2020 - 07/04/2020

muito discutido, sendo alvo de numerosos trabalhos. Perante este panorama, modelos preditivos estão sendo desenvolvidos de modo a identificar os principais fatores relacionados

Mesmo com a alta disponibilização de bases de dados governamentais, em sua grande maioria já pré-processadas, muitos estudos vêm utilizando bases referentes a hospitais, clínicas e centros médicos, com o objetivo de colher e avaliar uma maior variedade de informações de uma região específica. Além disso, pôde-se identificar que as estratégias propostas vêm sendo desenvolvidas e altamente treinadas, alcançando bons percentuais de acurácia. Entretanto, na grande maioria dos casos, os modelos preditivos não foram disponibilizados e tão pouco empregados em ambientes reais e nem disponibilizados para utilização.

6. Conclusão

Neste trabalho, foram expostos os resultados alcançados em uma RSL, visando identificar o cenário atual de trabalhos científicos sobre o desenvolvimento de soluções preditivas agravamento do quadro clínico de pacientes com COVID-19. Por meio dos resultados, foi constatado que: *i)* um conjunto de técnicas baseadas na aprendizagem de máquinas vêm sendo utilizadas durante a criação dos modelos, a citar: *Random Forest, SVM, SVM, XG-Boost e Neural Net MLP*; *ii)* as linguagens Python e R são as mais empregadas na análise dos dados e desenvolvimento das estratégias preditivas, e *iii)* dados provenientes de hospitais, centros médicos e clínicas locais vêm sendo selecionados com maior frequência para a composição dos modelos.

Em síntese, os estudos analisados podem contribuir sobre como a detecção e possibilidade de agravamento pode ser auxiliada por meio do uso de estratégias preditivas. Com isso, além de apresentar uma perspectiva sobre os trabalhos intrínsecos ao COVID-19, são destacadas as técnicas, ferramentas (linguagens de programação e *softwares*) e fonte de dados utilizadas, de modo a auxiliar possíveis futuros trabalhos na replicação ou aprimoramento dos modelos desenvolvidos. Como trabalhos futuros, pretende-se aumentar o número de fontes de busca utilizadas nesta pesquisa, de modo a identificar mais estudos que possuam relação com a temática abordada neste trabalho.

Referências

- Abdulaal, A., Patel, A., Charani, E., Denny, S., Mughal, N., and Moore, L. (2020). Prognostic modeling of covid-19 using artificial intelligence in the united kingdom: Model development and validation. *Journal of Medical Internet Research*.
- Assaf, D., Gutman, Y., Neuman, T., Segal, G., Amit, S., Gafen, S., Shilo, N., Epstein, A., Mor-Cohen, R., Biber, A., Rahav, G., Levy, I., and Tirosh, A. (2020). Utilization of machine-learning models to accurately predict the risk for critical covid-19. *Internal and Emergency Medicine*.
- Booth, A., Abels, E., and McCaffrey, P. (2020). Development of a prognostic model for mortality in covid-19 infection using machine learning. *Modern Pathology*.
- Brasil (2020). O que é covid? <https://coronavirus.saude.gov.br/sobre-a-doenca#o-que-e-covid>.
- Cheng, F., Joshi, H., Tandon, P., Freeman, R., Reich, D., Mazumdar, M., Kohli-Seth, R., Levin, M., Timsina, P., and Kia, A. (2020). Using machine learning to predict icu transfer in hospitalized covid-19 patients. *Journal of Clinical Medicine*.
- Das, A., Mishra, S., and Gopalan, S. (2020). Predicting covid-19 community mortality risk using machine learning and development of an online prognostic tool. *PeerJ*.
- Dresch, A., Lacerda, D. P., and Valle, J. A. (2015). *Design science research: método de pesquisa para avanço da ciência e tecnologia*. Bookman, Porto Alegre, 1th edition.
- Jamshidi, M., Lalbakhsh, A., Talla, J., Peroutka, Z., Hadjilooei, F., Lalbakhsh, P., Jamshidi, M., Spada, L., Mirmozafari, M., Dehghani, M., Sabet, A., Roshani, S., Roshani, S., Bayat-Makou, N., Mohamadzade, B., Malek, Z., Jamshidi, A., Kiani, S., Hashemi-Dezaki, H., and Mohyuddin, W. (2020). Artificial intelligence and covid-19: Deep learning approaches for diagnosis and treatment. *IEEE Access*, 8:109581–109595.
- Katzenschlager, S., Zimmer, A., Gottschalk, C., Grafeneder, J., Seitel, A., Maier-Hein, L., Benedetti, A., Weigand, J. L. M., McGrath, S., and Denking, C. (2020). Can we predict the severe course of covid-19 - a systematic review and meta-analysis of indicators of clinical outcome? *medRxiv*.
- Kim, H., Han, D., Kim, J., Kim, D., Ha, B., Seog, W., Lee, Y., Lim, D., Hong, S., Park, M., and Heo, J. (2020). An easy-to-use machine learning model to predict the prognosis of patients with COVID-19: Retrospective cohort study. *Journal of Medical Internet Research*, 22(11):e24225.
- Pan, P., Li, Y., Xiao, Y., Han, B., Su, L., Su, M., Li, Y., Zhang, S., Jiang, D., Chen, X., Zhou, F., Ma, L., Bao, P., and Xie, L. (2020). Prognostic assessment of COVID-19 in the intensive care unit by machine learning methods: Model development and validation. *Journal of Medical Internet Research*, 22(11):e23128.
- Parchure, P., Joshi, H., Dharmarajan, K., Freeman, R., Reich, D., Mazumdar, M., Timsina, P., and Kia, A. (2020). Development and validation of a machine learning-based prediction model for near-term in-hospital mortality among patients with COVID-19. *BMJ Supportive & Palliative Care*, pages bmjspcare–2020–002602.

- Shamsoddin, E. (2020). Can medical practitioners rely on prediction models for covid-19? a systematic review. *Evid Based Dent*, page 84–86.
- Shinde, G., Kalamkar, A., Mahalle, P., Dey, N., Chaki, J., and Hassanien, A. (2020). Forecasting models for coronavirus disease (covid-19): A survey of the state-of-the-art. *SN Computer Science*, 1:197.
- Vaid, A., Somani, S., Russak, A., Freitas, J., Chaudhry, F., Paranjpe, I., Johnson, K., Lee, S., Miotto, R., Richter, F., Zhao, S., Beckmann, N., Naik, N., Kia, A., Timsina, P., Lala, A., Paranjpe, M., Golden, E., Danieleto, M., Singh, M., Meyer, D., Reilly, P., Huckins, L., Kovatch, P., Finkelstein, J., Freeman, R., Argulian, E., Kasarskis, A., B.Percha, Aberg, J., Bagiella, E., Horowitz, C., Murphy, B., Nestler, E., Schadt, E., Cho, J., Cordon-Cardo, C., Fuster, V., Charney, D., Reich, D., Bottinger, E., Levin, M., Narula, J., Fayad, Z., Just, A., Charney, A., Nadkarni, G., and Glicksberg, B. (2020). Machine learning to predict mortality and critical events in a cohort of patients with COVID-19 in new york city: Model development and validation. *Journal of Medical Internet Research*, 22(11):e24018.
- WHO (2020). Coronavirus. https://www.who.int/health-topics/coronavirus#tab=tab_1.