

ÁRVORE DE DECISÃO PARA EXTRAÇÃO DE REGRAS EM FATORES DE EXAMES DE SANGUE DE PACIENTES COM COVID-19

Fabiula Augusta dos Santos (IC) e Leandro Augusto da Silva (Orientador)

Apoio: PIBIC Mackpesquisa

RESUMO

A pandemia de COVID-19 causada pelo vírus Sars-CoV-2 tem sido um objeto de estudo relevante devido ao seu impacto global. Com a geração de uma grande quantidade de dados, técnicas de Inteligência Artificial e Aprendizado de Máquina têm sido empregadas para o diagnóstico da doença. Neste contexto, esta pesquisa obteve sucesso no desenvolvimento de um sistema de Inteligência Artificial Explicável utilizando a metodologia de iterações sequenciais de mapas auto-organizáveis (Árvore de SOM) para identificar indícios de infecção por Sars-CoV-2 em pacientes com suspeita de COVID-19 através de seus hemogramas. A metodologia proposta consiste em treinar um modelo SOM inicial para visualizar e calcular as sobreposições de classes, selecionar os neurônios com baixa entropia para refinamento, subdividir o conjunto de dados original em novos subconjuntos e treinar um novo SOM em cada subconjunto. Com a segunda camada de refinamento, houve um aumento significativo no desempenho do modelo, melhorando as previsões em 21,6%. O modelo apresentou melhorias significativas com uma segunda camada de SOM, aumentando a acurácia em 21,6%. As variáveis mais relevantes foram MCH, MCV, Hemoglobina e MCHC. A exploração de faixas de valores dessas variáveis resultou em regras interpretáveis que contribuem para a compreensão das previsões. Essa se mostrou uma abordagem promissora que pode vir a auxiliar profissionais de saúde no diagnóstico de COVID-19, ao fornecer explicações transparentes para as tomadas de decisão.

Palavras-chave: COVID-19. Inteligência Artificial Explicável. Mapas Auto-organizáveis.

ABSTRACT

The COVID-19 pandemic caused by the Sars-CoV-2 virus has been a relevant object of study due to its global impact. With the generation of a large amount of data, Artificial Intelligence and Machine Learning techniques have been employed for disease diagnosis. In this context, this research succeeded in developing an Explainable Artificial Intelligence system using the Sequential Self-Organizing Maps (SOM) methodology to identify indications of Sars-CoV-2 infection in patients suspected of COVID-19 through their blood tests. The proposed methodology consists of training an initial SOM model to visualize and calculate class overlaps, selecting neurons with low entropy for refinement, subdividing the original dataset into new subsets, and training a new SOM on each subset. With the second layer of refinement, the model showed remarkable enhancements increasing improving its predictions by 21.6%. The most relevant variables were MCH, MCV, Hemoglobin, and MCHC. Exploring the ranges of these variable values resulted in interpretable rules that contribute to the understanding of predictions. This promising approach can assist healthcare professionals in COVID-19 diagnosis by providing transparent explanations for decision-making.

Keywords: COVID-19, Explainable Artificial Intelligence, Self-Organizing Maps (SOM).

1. INTRODUÇÃO

A pandemia causada pelo vírus *Sars_CoV_2* detectado pela primeira vez em dezembro de 2019 no distrito de Wuhan na China, tem sido um dos principais objetos de estudo da comunidade acadêmica devido ao seu impacto sem precedentes (SOUZA et al., 2021; BATISTA et al., 2020; VALLIM; SILVA, 2022). A quantidade de informação gerada pela alta exponencial de internações clínicas de pessoas com COVID-19 ou com suspeita, gerou grandes bases de dados permitindo estudos mais aprofundados acerca desta doença.

Como alguns exames e testes não são plenamente acessíveis e, mesmo os disponíveis, estão sujeitos a problemas relativos à baixa dos estoques ou adequados procedimentos de armazenagem e transporte, cujas eventuais falhas comprometem lotes inteiros de testes. Assim, os dados gerados na investigação do diagnóstico como o exame de sangue é possível ter, como alternativa, abordagens como Inteligência Artificial e Aprendizagem de Máquina (SOUZA et al., 2021; BATISTA et al., 2020; VALLIM; SILVA, 2022). Deste modo, pode-se gerar modelos que permitem o agrupamento e classificação diagnóstica de pacientes conforme os dados providos de exames rotineiros e baratos como é o caso do hemograma. Com isso é possível diagnosticar com mais agilidade os quadros infecto contagiosos, como visto em diferentes estudos que buscaram utilizar algoritmos de aprendizagem de máquina para identificar pacientes com COVID-19. Alguns destes algoritmos foram: máquina de suporte vetor (SVM), Naive Bayes, redes neurais, random forests, regressão logística, mapas auto-organizáveis (SOM), entre outros (SOUZA et al., 2021; BATISTA et al., 2020; VALLIM; SILVA, 2022).

Embora algumas dessas abordagens tenham atingido índices de acurácia média superior a 80%, se destacam os trabalhos de Souza et al. (2021) e Vallim e Silva (2022) que buscaram, além de identificar a infecção, explicar os diagnósticos identificando as principais variáveis para o resultado obtido.

A abordagem apresentada por Vallim e Silva (2022) e Souza et al (2021) trazem evidências de que é possível identificar, por meio de hemogramas, se pacientes com suspeita de Covid-19 de fato estão infectados com o vírus. Mas, além disso, tais trabalhos visam identificar as principais variáveis do hemograma para a realização do diagnóstico. Vallim e Silva (2022) ainda vão além, buscando identificar os intervalos de valores dessas variáveis, criando regras para a definição do diagnóstico. Mas, apesar de ser de fácil interpretação, o algoritmo utilizado para a criação de regras, Classificação Baseada em Regras de Associação apresenta pouco apelo visual que pode ser de grande utilidade para o usuário final. Dessa forma, um modelo com o desenvolvido neste trabalho, consistindo de

uma Árvore de SOM tem como resultado representações visuais de fácil interpretação e que podem ser implementadas em sistemas computacionais, como o caso de *dashboards*.

Esse tipo de abordagem – visualização de regras usadas para tomada de decisão - é relevante, pois quando modelos de aprendizagem de máquina são implementados na área da saúde é essencial que o modelo tenha um equilíbrio entre desempenho e interpretabilidade, em vista que os profissionais da saúde buscam entender os motivos que geram determinada saída no sistema (LAMY et al., 2019; WU et al., 2021; HOLZINGER, 2018). Dessa forma, é importante que os modelos sejam capazes de apresentar alta acurácia em suas predições, mas também é necessário que sejam interpretáveis pelos seus usuários. (ADADI; BERRADA, 2018; LAMY et al.).

Com base no que foi apresentado, o estudo aqui desenvolvido buscou responder a seguinte questão: É possível desenvolver um sistema de Inteligência Artificial Explicável com utilizando uma metodologia de iterações sequenciais de mapas auto-organizáveis (aqui chamada de “Árvore de SOM”) que seja capaz de apresentar altos índices de desempenho e identificar as principais variáveis de um hemograma de pacientes que procuraram o hospital com suspeita de COVID, bem como os intervalos de valores das variáveis que se relacionam com os diagnósticos?

A partir desta pergunta, este trabalho buscou, como objetivo geral, desenvolver uma abordagem iterativa e sequencial de Mapas Auto-organizáveis visando compreender relações entre as variáveis medidas em hemogramas simples e o diagnóstico de COVID-19.

Descrito o objetivo geral, os objetivos específicos do estudo consistiram em explorar o uso de Mapas Auto-organizáveis para visualizar eventuais sobreposições de classes e suas concentrações; Calcular as sobreposições de cada neurônio gerado pelo modelo SOM; Implementar, nos neurônios que apresentarem sobreposição de classes, um Mapas Auto-organizáveis incremental para extração de regras e identificar as principais variáveis para a realização do diagnóstico de COVID-19; e ser capaz de interpretar o modelo e identificar as faixas de valores das variáveis selecionadas como mais relevantes.

2. REFERENCIAL TEÓRICO

A inteligência artificial e aprendizagem de máquina são campos de estudo que vem adquirindo cada vez mais importância em diversas áreas da sociedade. Apesar disso, o conceito de inteligência artificial não é um conceito relativamente novo, tendo surgido nos anos 1950. Segundo Chollet (2018), Inteligência Artificial pode ser definida como o esforço

de automatizar tarefas intelectuais comumente executadas por seres humanos. Para Russel e Norvig (2010) a definição de Inteligência Artificial é a ciência que busca alcançar os melhores resultados por meio de desenvolvimento de sistemas computacionais autônomos.

Antes da Aprendizagem de Máquina a Inteligência Artificial dependia da criação de regras que fossem capazes de suficientemente abranger todas as situações que poderiam ocorrer, de tal forma que as equipes de desenvolvimento desses sistemas apresentassem extraordinária capacidade analítica de antecipação de cenários, inserindo os dados e suas respostas esperadas na codificação para que assim o sistema tivesse a capacidade de gerar regras relacionando essas informações (CHOLLET, 2018).

Mesmo se mostrando adequada na solução de problemas lógicos era necessário outro conjunto de técnicas para casos mais complexos, surgindo assim a Aprendizagem de Máquina, uma abordagem que visa explorar e descobrir a possibilidade de computadores realizarem tarefas de formas autônomas sem necessitarem de intermédio humano (CHOLLET, 2018; RUSSEL; NORVIG, 2010).

Ao invés de utilizar regras previamente estipuladas pelo usuário, para então haver processamento de dados, a Aprendizagem de máquina utiliza a descoberta de conhecimento e aprendizado diretamente por meio dos dados de entrada, sendo assim é possível definir a Aprendizagem de Máquina como a ciência que visa o desenvolvimento de modelos computacionais com capacidade de aprendizagem justamente a partir dos dados (CHOLLET, 2018;).

Russel e Norvig (2010) diferenciam os processos a partir do paradigma de *feedback* adquirido pelos algoritmos. Desse modo é possível destacar que os autores identificam três principais paradigmas de aprendizado, sendo eles: Aprendizado Não Supervisionado, Aprendizado Supervisionado e Aprendizado Por Reforço.

O Aprendizado não supervisionado se caracteriza pelo descobrimento de aprendizagem de padrões pelo algoritmo sem rotulações, de tal forma que o modelo busca informações nos dados de entrada sem a indicação de rótulos ou valores esperados que correspondam aos dados de entrada (CHOLLET, 2018; RUSSEL; NORVIG, 2010).

No caso do paradigma de aprendizagem supervisionada, o modelo é “alimentado” com dados de entrada a informação correspondente de saída, de tal forma que o algoritmo busca aprender a relação entre esses dois conjuntos. (CHOLLET, 2018; RUSSEL; NORVIG, 2010).

O aprendizado por reforço segue uma lógica ligeiramente diferente dos dois paradigmas anteriores. O algoritmo recebe informações sobre um ambiente no qual ele é inserido e busca mapear eventuais ações ou escolhas que podem vir a ser executadas, sendo retroalimentado por reforços que recompensam ou punem as decisões anteriores, de tal forma que o modelo se ajusta, buscando encontrar soluções mais otimizadas baseado por esses reforços obtidos após cada rodada. (CHOLLET, 2018; RUSSEL; NORVIG, 2010).

Apesar de apresentar ótimos resultados e apresentar diferentes abordagens que podem ser aplicadas em diferentes contextos, muitos dos algoritmos de aprendizagem de máquina que apresentam os melhores resultados, são os algoritmos “caixa-preta” que apresentam uma estrutura interna desconhecida ou pouco compreensível para os usuários. Esse fato, segundo Adadi e Berrada (2018) limita as análises desses modelos apenas às relações entre conjuntos de dados de entrada e as respectivas saídas fornecidas.

Esse fato não possibilita a interpretação adequada de predições, o que pode se tornar um problema quando os resultados não são intuitivos, por exemplo. Para que sejam desenvolvidas melhorias em um sistema, é necessário compreender o processo de tomada de decisão do modelo de Inteligência Artificial, facilitando assim o processo de desenvolvimento (ADADI; BERRADA, 2018; HOLZINGER, 2018).

Em algumas áreas, como a medicina diagnóstica, por exemplo, que convive com tomadas de decisões críticas de forma rotineira, é necessário que os especialistas estejam embasados em seus processos de tomada de decisão de tal forma que seja possível entender o que levou ao diagnóstico fornecido. Desse modo, é essencial que recomendações fornecidas por modelos de Inteligência Artificial sejam interpretáveis, e a existência de modelos caixa-preta limita a adoção dessa tecnologia nessas áreas. (ADADI; BERRADA, 2018; DOSILOVIC; BRCIC; HLUPIC, 2018).

Para lidar com essa questão, surge a chamada Inteligência Artificial Explicável, que é definida por Adadi e Berrada (2018) como a ciência de tornar resultados de algoritmos de Inteligência Artificial mais compreensíveis para os seres humanos.

O objetivo da Inteligência Artificial Explicável é produzir modelos que, sem comprometer o desempenho dos modelos, explicar recomendações e predições para os seres humanos.

Então é importante destacar que a discussão sobre a explicação na tomada de decisão em algoritmos de Inteligência Artificial, geralmente significa fornecer justificativas para as saídas do modelo, e não necessariamente a descoberta e explanação de cálculos

internos do algoritmo utilizado ou mesmo a lógica utilizada (ADADI; BERRADA, 2018; HOLZINGER 2018).

3. PROCEDIMENTOS METODOLÓGICOS

Nesta seção são listados e descritos os procedimentos metodológicos que foram adotados para o desenvolvimento do estudo desenvolvido.

3.1 COLETA E FILTRAGEM DE DADOS

Os dados coletados para a utilização neste estudo consistem em bases de dados de diferentes centros hospitalares do estado de São Paulo. Estes dados estão disponíveis no repositório COVID-19 *Data Sharing/BR* da FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) em cooperação com a Universidade de São Paulo (USP) com o objetivo de disponibilizar dados relacionadas à COVID-19 que possam contribuir para pesquisas (FAPESP, 2020).

Os dados disponíveis neste repositório representam dados de 2178 hemogramas simples de pacientes com suspeita de COVID-19 nestes centros hospitalares, bem como seus respectivos testes PCR indicando a presença ou não de infecção pelo vírus *Sars_CoV_2*. Ao observar o resultado dos exames, verifica-se que 1089 testaram positivo para Covid-19 e 1089 testaram negativo (FAPESP, 2020).

Além de conter o resultado do exame de PCR, a base contém outros 13 atributos, representantes do hemograma mencionado. Os atributos e suas respectivas descrições, de acordo com Alves et al. (2021) são:

Hematocrit - O hematócrito, também conhecido como Hct ou Ht, é a medida que indica a porcentagem de células vermelhas, no volume total de sangue;

Hemoglobin - Mede a quantidade de hemoglobina no sangue. É uma proteína presente nos glóbulos vermelhos. Tem como função o transporte de oxigênio pelo organismo. Valores de referência da hemoglobina varia de acordo com o sexo biológico do paciente;

Platelets - Mede a quantidade de plaquetas no sangue. Células que são responsáveis pelo processo de coagulação sanguínea;

Red.Blood.Cells - Contagem do número de hemácias (células vermelhas) no sangue;

Lymphocytes.Absolute - Contagem absoluta de linfócitos presentes no sangue. É um tipo de leucócito presente no sangue com importante papel no sistema imunológico. Recebeu esse nome por ser o principal tipo de célula encontrada na linfa;

Mean.corpuscular.hemoglobin.concentraion.MHCH - É o índice que avalia a concentração de hemoglobina presente nas hemácias;

Leukocytes - A contagem de Leucócitos, os glóbulos brancos, no sangue. A principal função é a defesa do organismo contra agentes externos;

Basophils.Absolute - Contagem absoluta de basófilos. São um tipo específico de glóbulo branco que atua na detecção de agentes externos e no reparo de feridas;

Mean.corpuscular.hemoglobin.MCH - Mede a quantidade média de hemoglobina presente nas hemácias. Pode ser usada para detectar, por exemplo anemia;

Eosinophils.Absolute - Contagem absoluta de eosinófilos. São um tipo de célula de defesa que, durante infecções ficam presentes no sangue em elevadas concentrações;

Mean.corpuscular.volume.MCV - É um índice que indica a média do tamanho das hemácias, que são as células vermelhas do sangue;

Monocytes.Absolute - Contagem absoluta de monócitos no sangue. São um tipo de glóbulo branco que, entre outras tarefas, ajuda no combate de infecções, remover tecidos mortos e regular a imunidade contra substâncias estranhas; e

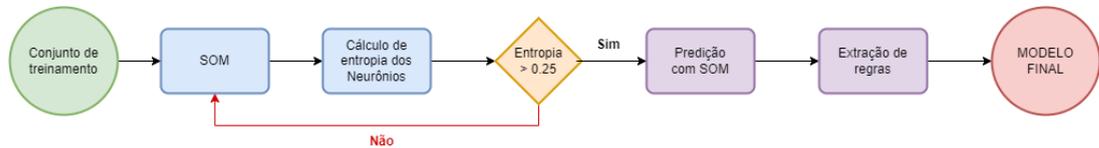
Red.blood.cell.distribution.width.RDW - Mede a amplitude de distribuição de todos glóbulos vermelhos, que é, basicamente, a variação de tamanho entre as hemácias (anisocitose);

3.2 CONSTRUÇÃO DO MODELO

O método proposto neste trabalho consistiu em desenvolver um modelo de Inteligência Artificial Explicável capaz de identificar, por meio de hemogramas, indícios da infecção por SARS-CoV-2 no organismo de pacientes com suspeita de Covid-19

Para atingir este objetivo, este modelo é composto pela implementação sequencial de diferentes camadas do Mapa Auto-Organizável conforme o diagrama demonstrado na figura 1.

Figura 1: Diagrama descrevendo o fluxo do modelo proposto.



Fonte: Própria Autora

Na figura 1 é possível observar que o modelo proposto consiste em 3 etapas distintas, sendo estas, uma dedicada ao desenvolvimento do mapa SOM inicial, uma etapa dedicada à análise de sobreposição de classes calculando a entropia de Shannon em cada neurônio gerado pelo mapa e a última dedicada à uma tomada de decisão, aqueles neurônios que apresentam entropia abaixo do limite definido são submetidos à extração das regras enquanto os demais se tornam subconjuntos e retornam, cada um para uma nova sequência, treinando-se um novo SOM e repetindo as etapas.

3.4 ANÁLISE DE DESEMPENHO DO MODELO

Baseado, principalmente, nos índices de acurácia, o desempenho do modelo construído foi avaliado com relação ao sucesso na avaliação diagnóstica de COVID-19 dos pacientes.

3.5 IDENTIFICAR VARIÁVEIS MAIS RELEVANTES E EXTRAIR REGRAS

Foram identificadas, a partir do modelo construído, as variáveis mais relevantes na obtenção de diagnósticos de COVID-19, bem como extraídos os intervalos de valores dessas variáveis que se relacionam com os resultados do exame PCR.

4. RESULTADOS OBTIDOS

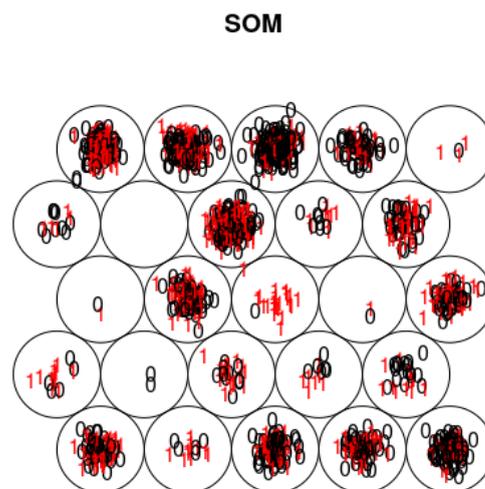
Nesta seção são listados e descritos os experimentos realizados pelo modelo descrito na seção anterior, bem como seus respectivos principais resultados obtidos.

4.1 IMPLEMENTAÇÃO DO MODELO

O experimento aqui apresentado consiste em treinar o modelo proposto anteriormente, separando a base de forma aleatória em 75% dos dados para treinamento e 25% dos dados para teste.

Inicialmente foi desenvolvido o modelo SOM com as dimensões 5x5 visando visualizar e calcular as sobreposições entre as classes. O resultado pode ser visto na Figura 2.

Figura 2: Mapa gerado durante o processo de treinamento.



Fonte: Própria Autora

Os elementos apresentados no mapa são os objetos alocados nos neurônios correspondentes após o processo de treinamento do algoritmo.

O algoritmo SOM analisa as variáveis sem informação de classe de maneira não supervisionada, organizando os dados com base em suas similaridades. Isso leva à alocação de objetos semelhantes próximos ao longo do mapa.

Dessa forma, é possível observar, analisando a figura 2, que alguns neurônios mostram uma concentração intensa de objetos sobrepostos, enquanto outros possuem menos objetos e menor sobreposição.

Nos neurônios com maior número de objetos, a sobreposição de classes é mais evidente. Para quantificar o grau de impureza desses conjuntos gerados, foi necessário calcular a Entropia de Shannon.

Os valores da Entropia de Shannon para cada neurônio podem ser observados na tabela 1.

Tabela 1: Tabela com as entropias de Shannon calculadas para cada neurônio do SOM.

Neurônio	Entropia	Neurônio	Entropia
1	0.6890092	14	0.6931472
2	0.6365142	15	0.6924612
3	0.6655035	16	0.6517566
4	0.6869616	17	0
5	0.6730117	18	0.6788449
6	0.6057975	19	0.6931472
7	0	20	0.6751426
8	0.6591525	21	0.6914659
9	0.6901857	22	0.6685396
10	0.6829081	23	0.5750398
11	0.6931472	24	0.6588827
12	0.6876496	25	0.5623351
13	0.1679441		

Fonte: Própria Autora

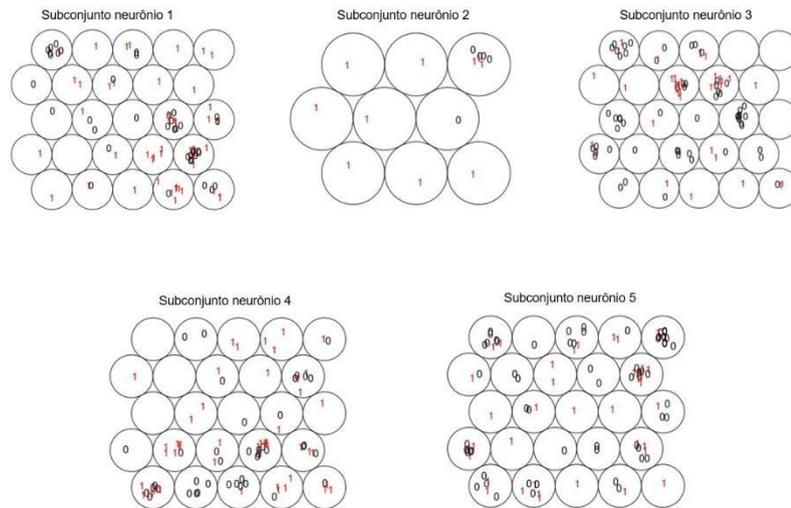
Seguindo o critério apontado anteriormente, o objetivo ao definir o valor pré-determinado de 0,25 foi separar os neurônios em dois grupos: aqueles que exibem sobreposição de classes e aqueles que não apresentam essa sobreposição. Como observado na tabela 1, somente os neurônios 7, 13 e 17 apresentaram entropia inferior ao limite estabelecido, forçando com que os outros 22 neurônios fossem submetidos à próxima etapa do modelo.

Esses valores de entropia resultaram na subdivisão do conjunto original de dados em 22 subconjuntos que foram utilizados como conjunto de entrada para novos modelos SOM com o intuito amplificar o desempenho do modelo, bem como extrair melhor as regras de determinação de diagnóstico.

Os mapas gerados para cada subconjunto derivado de cada neurônio do SOM original podem ser vistos nas figuras 3, 4, 5, 6, e 7.

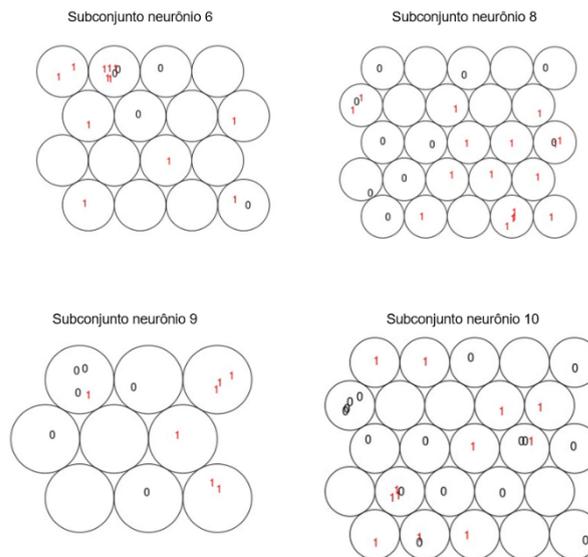
Como é possível observar, a distribuição de observações em cada neurônio se deu de forma bastante heterogênea. Dessa forma, para cada subconjunto foi necessário realizar um ajuste nas dimensões do reticulado do mapa de tal modo que alguns mapas tiveram dimensões bastante reduzidas para a realização do treinamento do modelo.

Figura 3: Mapas gerados durante o processo de treinamento do SOM com os subconjuntos originários dos neurônios 1, 2, 3, 4 e 5.



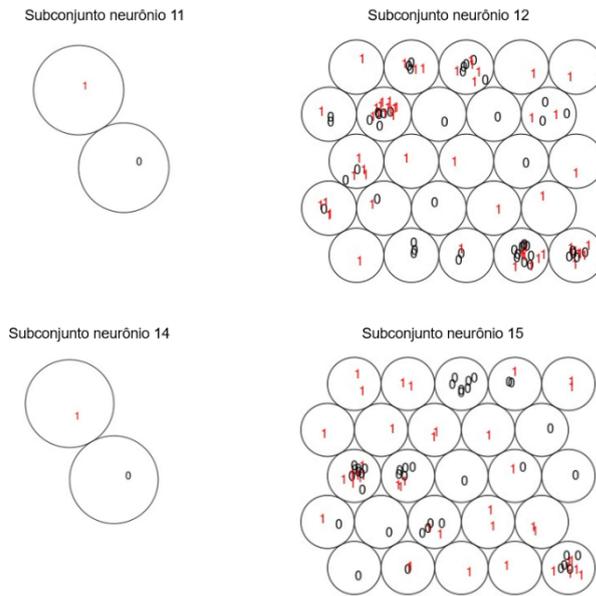
Fonte: Própria Autora

Figura 4: Mapas gerados durante o processo de treinamento do SOM com os subconjuntos originários dos neurônios 6, 8, 9 e 10.



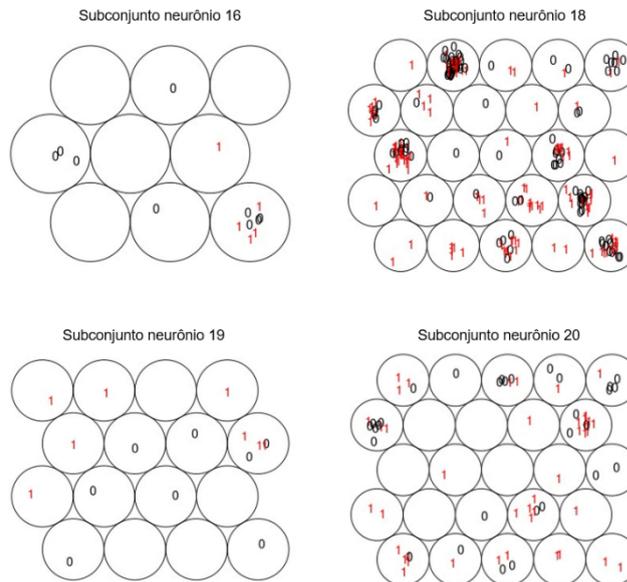
Fonte Própria Autora

Figura 5: Mapas gerados durante o processo de treinamento do SOM com os subconjuntos originários dos neurônios 11, 12, 14 e 15.



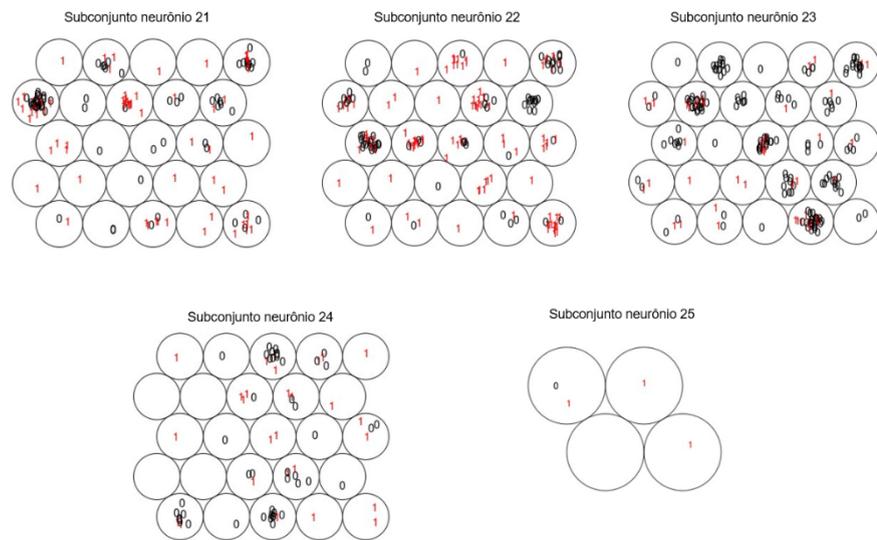
Fonte: Própria Autora

Figura 6: Mapas gerados durante o processo de treinamento do SOM com os subconjuntos originários dos neurônios 16, 18, 19 e 20.



Fonte: Própria Autora

Figura 7: Mapas gerados durante o processo de treinamento do SOM com os subconjuntos originários dos neurônios 21, 22, 23, 24 e 25.



Fonte: Própria Autora

A análise dos subconjuntos apresentados nas figuras 3, 4, 5, 6, e 7 deixa evidente a menor sobreposição das classes em relação ao mapa original. Essa informação é fundamentada pelas análises das entropias de Shannon desses novos mapas.

Enquanto a média de entropia dos 25 neurônios do mapa original é de aproximadamente 0,59 com um desvio padrão de aproximadamente 0,21, a média de entropia dos neurônios da segunda iteração do modelo é de 0,22 com um desvio padrão de 0,3. A média foi reduzida em 62% e se aproximou muito do limite pré determinado de 0,15.

A melhoria nos resultados também é apoiada pelo desempenho observado quando comparadas as previsões realizadas diretamente na primeira etapa do modelo em contraponto ao refinamento da segunda iteração.

Ao realizar a predição com o modelo gerado pelo primeiro mapa, sem refinamento, os resultados podem ser observados na matriz de confusão apresentada na figura 8.

Figura 8: Matriz de Confusão resultante do treinamento do SOM sem nenhum refinamento.

		Predição	
		0	1
Real	0	356	457
	1	185	635

Fonte: Própria Autora

Na figura 8, é possível observar as previsões resultantes do treinamento do SOM em comparação com os resultados reais. Na diagonal principal da matriz estão registradas

as quantidades de classificações corretas que o modelo realizou, sendo 356 classificações negativas para COVID-19 e 635 positivas. De maneira análoga, o modelo classificou 457 pacientes com COVID quando o PCR não forneceu este diagnóstico, resultando em um falso positivo e também classificou 185 pacientes de forma negativa em contraponto ao resultado do PCR, resultando em falso negativo.

Já os resultados obtidos com o refinamento da segunda iteração podem ser visualizados na figura 9.

Figura 9: Matriz de Confusão resultante do treinamento com a subdivisão e a segunda camada de SOM.

		Predição	
		0	1
Real	0	642	171
	1	257	563

Fonte: Própria Autora

Na figura 9, é possível observar as predições resultantes do treinamento do SOM com uma segunda camada de refinamento, em comparação com os resultados do PCR. Da mesma forma que na figura 8, a diagonal principal da matriz apresenta as quantidades de classificações corretas que o modelo realizou, sendo 642 classificações negativas para COVID-19 e 563 positivas. De maneira análoga, o modelo classificou 171 pacientes com COVID quando o PCR não forneceu este diagnóstico, resultando em um falso positivo e também classificou 257 pacientes de forma negativa em contraponto ao resultado do PCR, resultando em falso negativo.

Os resultados comparados de acurácia apresentados pelas matrizes de confusão das figuras 8 e 9 mostram que o refinamento da segunda iteração melhorou a predição final em 21,6% com uma mudança de 60,7% para 73,8%, evidenciando que a metodologia se mostra eficiente na busca por melhores resultados.

Para fins de comparação de resultados, os índices obtidos por de Souza et al (2021) e Vallim e Silva (2022), apresentam-se em um patamar semelhante. O resultado obtido em cada trabalho foi de, respectivamente, 83% e 74,5%. No caso do trabalho de Souza et al (2021), apesar da alta acurácia, houve um desbalanceamento de resultados, pois o índice de sensibilidade foi de 56%.

Ambos os trabalhos apresentaram índices de acurácia ligeiramente melhores, mas, a árvore de SOM desenvolvida neste trabalho contou com apenas uma iteração de

refinamento, evidenciando um potencial real de resultados robustos e maior balanceamento entre sensibilidade e especificidade.

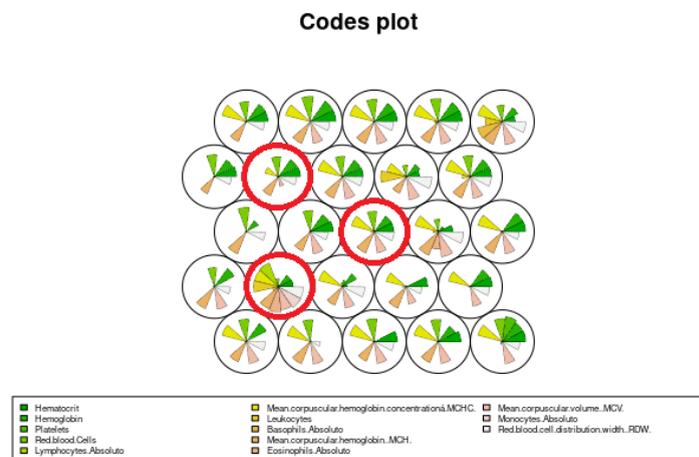
Importante notar que, além de buscar por melhores índices de acurácia, o modelo também se mostrou eficiente na explicabilidade das predições.

4.2 EXTRAÇÃO DE REGRAS

Para realizar a tarefa de extração das regras e explicação das predições realizadas pelo modelo, foram levantadas as variáveis mais relevantes de acordo com os mapas gerados.

O mapa que destaca a importância de cada variável na alocação de observações em cada neurônio podem ser vistos nas figuras 10, 11, 12, 13, 14 e 15, sendo que, na figura 10 está o mapa SOM original com o devido destaque aos neurônios de baixa entropia e que, conseqüentemente, foram utilizados na classificação do modelo.

Figura 10: Variáveis mais relevantes para alocação de neurônios no SOM original e destaque para os neurônios de entropia inferior a 0,25.

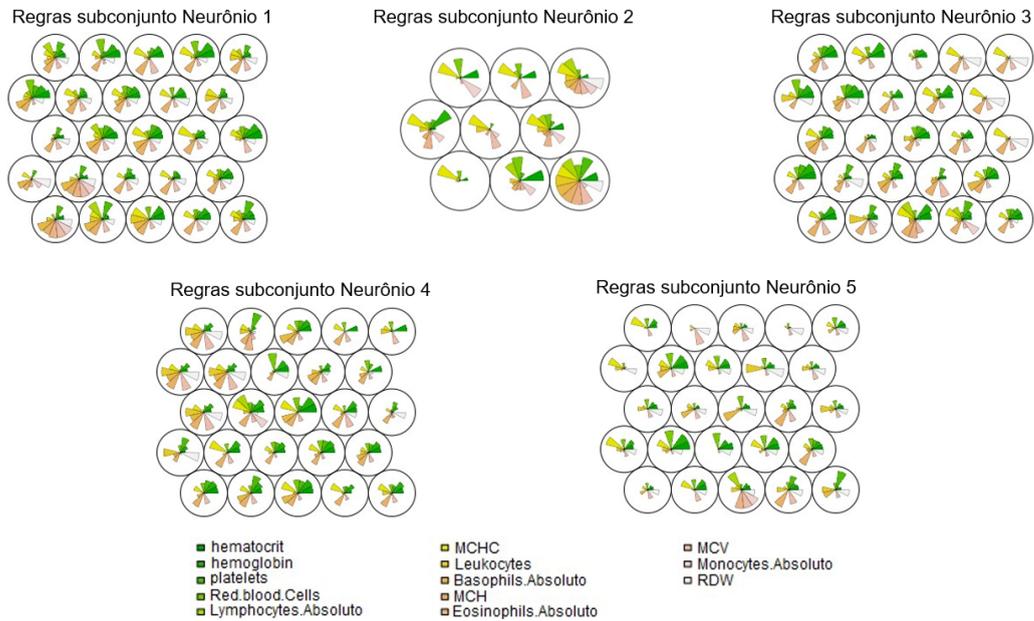


Fonte: Própria Autora

As figuras 10, 11, 12, 13, 14 e 15 mostram, dentro de cada neurônio de cada mapa SOM representado, gráficos que destacam o nível de importância de cada uma das 13 variáveis na alocação de observações em cada neurônio durante o período de testes. Um levantamento estatístico de quais as variáveis aparecem mais frequentemente entre as principais demonstra uma consistência na presença de 4 variáveis principais. Dos 420 neurônios analisados (todos os neurônios dos 22 mapas SOM gerados na segunda camada

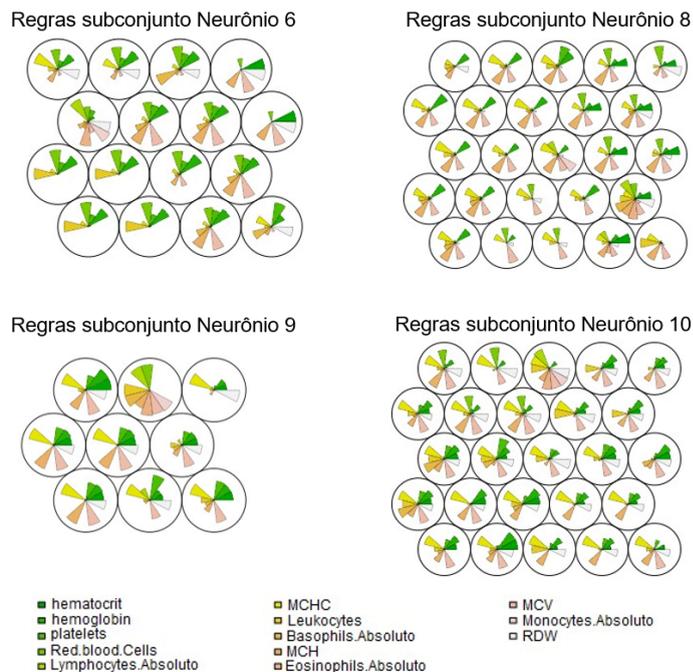
mais os 3 neurônios de baixa entropia no SOM original), as variáveis MCH, MCV, Hemoglobina e MCHC se destacaram acima das demais com frequências altas.

Figura 11: Variáveis mais relevantes para alocação de neurônios nos mapas SOM criados com os subconjuntos 1, 2, 3, 4 e 5.



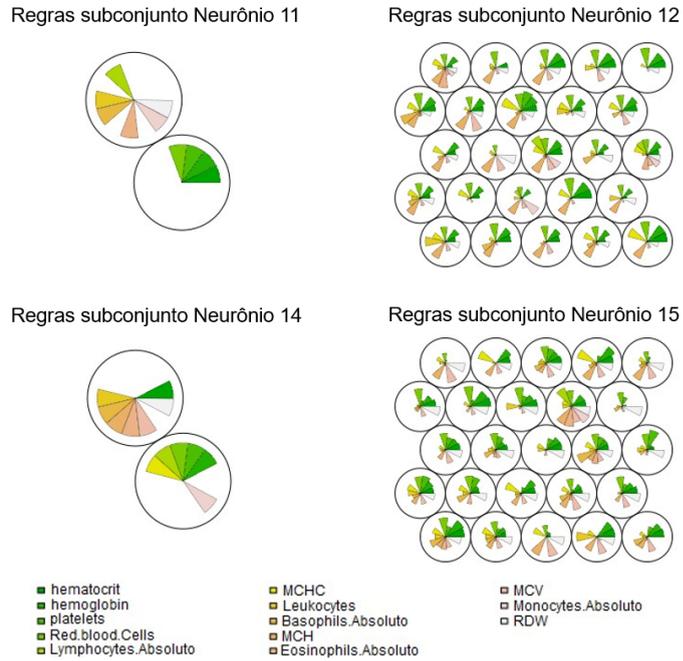
Fonte: Própria Autora

Figura 12: Variáveis mais relevantes para alocação de neurônios nos mapas SOM criados com os subconjuntos 6, 8, 9 e 10.



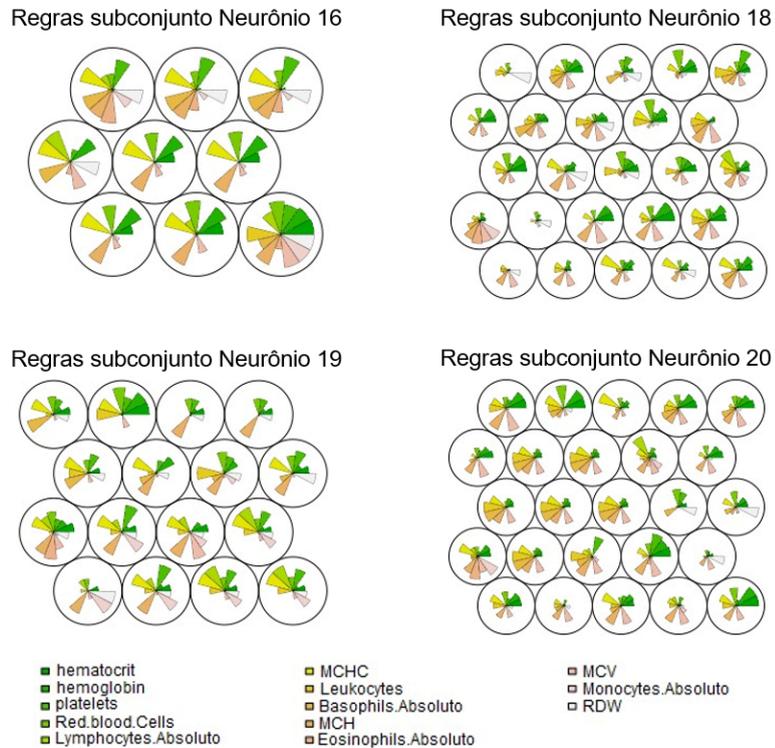
Fonte: Própria Autora.

Figura 13: Variáveis mais relevantes para alocação de neurônios nos mapas SOM criados com os subconjuntos 11, 12, 14 e 15.



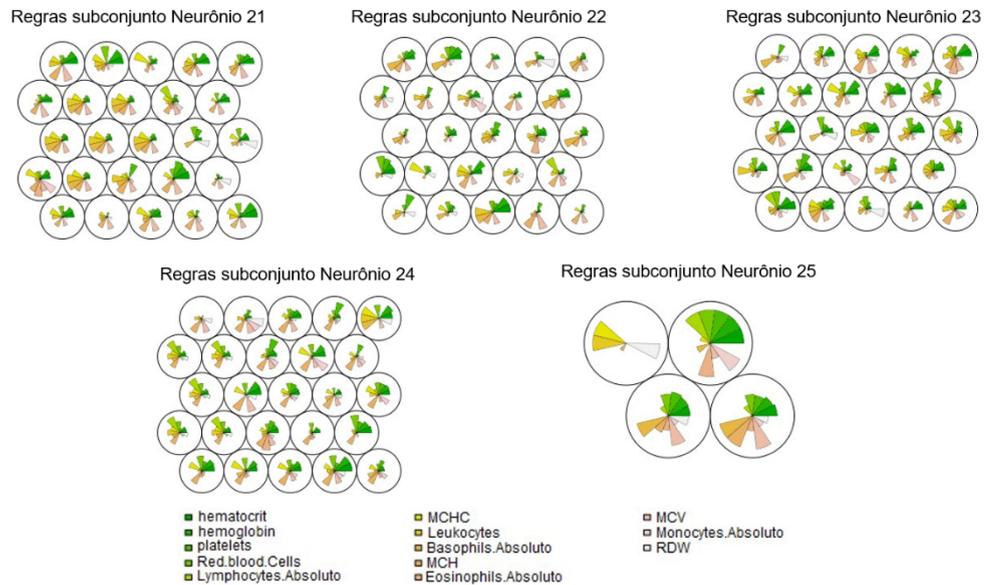
Fonte: Própria Autora.

Figura 14: Variáveis mais relevantes para alocação de neurônios nos mapas SOM criados com os subconjuntos 16, 18, 19 e 20.



Fonte: Própria Autora.

Figura 15: Variáveis mais relevantes para alocação de neurônios nos mapas SOM criados com os subconjuntos 21, 22, 23, 24 e 25.



Fonte: Própria Autora.

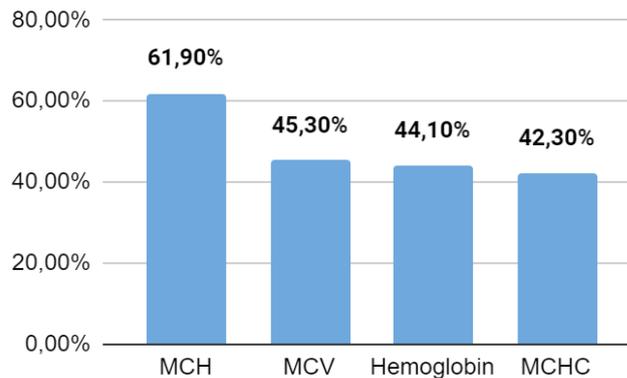
As frequências absolutas e relativas podem ser vistas na tabela 2 e as frequências relativas podem ser analisadas visualmente na figura 16.

Tabela 2: Frequências relativa e absoluta das variáveis mais relevantes para o diagnóstico de COVID-19 em pacientes sintomáticos.

Variável	Freq. Relativa	Freq. Absoluta
MCH	61,90%	260
MCV	45,30%	190
Hemoglobin	44,10%	185
MCHC	42,30%	178

Fonte: Própria Autora.

Figura 16: Variáveis relevantes mais frequentes em todos os neurônios do levantamento.



Fonte: Própria Autora.

Para fins de comparação, o trabalho apresentado por Vallim e Silva (2022) extraiu regras em dois cenários diferentes: utilizou CBA para extrair regras em áreas de sobreposição de classes e utilizou o mesmo método deste trabalho para extrair regras em áreas de pouca sobreposição de classes.

É importante notar que as variáveis relevantes extraídas por Vallim e Silva (2022) foram diferentes em ambos os cenários. Utilizando o CBA houve completa divergência das variáveis, mas no cenário em que o método foi o mesmo, o SOM identificou 4 variáveis relevantes, sendo 3 convergentes com o deste trabalho: MCV, MCHC e MCH. Evidenciando que as variáveis aqui identificadas podem mostrar, de fato, um caminho para o diagnóstico de COVID-19.

As frequências absolutas e relativas apresentadas na tabela 2 e na figura 16 servem de referência para a extração eficiente de regras, como são as variáveis que, consistentemente, aparecem como as mais relevantes, entender suas faixas de variação de valores se torna essencial para a explicabilidade dos resultados das classificações.

A partir disso, a análise para extração da assinatura das variáveis foi feita baseada na mineração de dados e busca de faixas de oscilação e eventual identificação de *outliers* dentro dos conjuntos formados em cada neurônio. O resultado dessa mineração das regras geradas pelo SOM, os valores podem ser observados na tabela 3.

Tabela 3: Regras mineradas da base de dados.

Variável	Regra	Diagnóstico SARS-CoV-2
MCH	0 a 25	Negativo
	25 a 279	Positivo
	281 a 306	Negativo
	acima de 399	Positivo
MCV	Até 876	Positivo
	876 a 1204	Negativo
	acima de 1204	Positivo
Hemoglobin	0 a 7	Negativo
	122 a 185	Área de Indefinição
	acima de 185	Positivo
MCHC	0 a 31	Negativo
	31 a 325	Positivo
	324 a 342	Área de Indefinição
	369 a 374	Negativo

Fonte: Própria Autora.

As regras extraídas na tabela 3 convergem para faixas de valores semelhantes, especialmente MCV e MCH. Este resultado evidencia que, mesmo com metodologias diferentes, há um padrão para a identificação das infecções.

5. CONCLUSÃO E PRÓXIMOS PASSOS

Os resultados apontam um ganho de desempenho considerável entre a implementação do modelo proposto e o SOM implementado sem nenhuma etapa de refinamento. Também foi possível avaliar que as variáveis apontadas como mais relevantes para a realização de diagnósticos de contágio por *Sars-CoV-2* são MCH, MCV, Hemoglobina e MCHC, o que demonstra consistência com outros estudos.

Os resultados obtidos apresentam índices de desempenho próximos a outros trabalhos, mas com um potencial de ganho de desempenho superior, dado o fato de que apenas uma iteração do refinamento foi implementada.

As variáveis destacadas como relevantes também apontam uma convergência com outros trabalhos, bem como as faixas de valores extraídas para a identificação do diagnóstico.

Para estudos futuros, é recomendado avaliar a implementação de novas camadas de SOM, medindo ganho de desempenho e consistência das regras extraídas, além de compreender a quantidade ideal de novas camadas.

6. REFERÊNCIAS BIBLIOGRÁFICAS

- ADADI, Amina; BERRADA, Mohammed. Peeking inside the black-box: a survey on explainable artificial intelligence (XAI). **IEEE access**, v. 6, p. 52138-52160, 2018.
- CHOLLET, Francois. **Deep Learning With Python**. Shelter Island: Manning. [S.l.]: Manning, 2018.
- DE MORAES BATISTA, Andre Filipe et al. COVID-19 diagnosis prediction in emergency care patients: a machine learning approach. **MedRxiv**, 2020.
- DOŠILOVIĆ, Filip Karlo; BRČIĆ, Mario; HLUPIĆ, Nikica. Explainable artificial intelligence: A survey. In: **2018 41st International convention on information and communication technology, electronics and microelectronics (MIPRO)**. IEEE, 2018. p. 0210-0215.
- FAPESP. FAPESP COVID-19 Data Sharing/BR. 2020. Disponível em: <<https://repositoriodatasharingfapesp.uspdigital.usp.br/>>, Acessado em: 05 de abril de 2022.
- HOLZINGER, Andreas. From machine learning to explainable AI. In: **2018 world symposium on digital intelligence for systems and machines (DISA)**. IEEE, 2018. p. 55-66.
- RUSSEL, S.; NORVIG, P. Artificial Intelligence A Modern Approach. 3. ed. [S.l.]: Pearson, 2010.
- SOUZA, Alexandra A. de et al. Simple hemogram to support the decision-making of COVID-19 diagnosis using clusters analysis with self-organizing maps neural network. **Soft Computing**, p. 1-12, 2021.
- VALLIM, M. V. B. A; Inteligência Artificial Explicável Aplicada a Hemogramas como suporte à tomada de decisão em diagnósticos de COVID-19. *Orientador: Leandro Augusto da Silva*. 2022. 112. Dissertação (Mestrado) - Engenharia Elétrica e Computação, Mackenzie, São Paulo, 2022.
- WU, Han et al. Interpretable machine learning for covid-19: An empirical study on severity prediction task. **IEEE Transactions on Artificial Intelligence**, 2021.

Contatos: fabiula.augustasantos@gmail.com (aluna) e leandroaugusto.silva@mackenzie.br (orientador).