

## Predição da taxa de ocupação das UTIs nas regiões brasileiras considerando a vacinação

Predicting bed occupancy rate of ICUs in Brazilian regions considering vaccination

La predicción de la tasa de ocupación de las UCIs en las regiones brasileñas considerando la vacunación

Gabriel Caldas Barros e Sá<sup>1</sup>, Augusto Vinícius da Silva<sup>2</sup>, Náthalee Cavalcanti de Almeida Lima<sup>3</sup>,  
Samara Martins Nascimento<sup>3</sup>

### RESUMO

**Descritores:** Leitos de UTI; *Machine Learning*; COVID-19

**Objetivos:** Propor um modelo de *Machine Learning* para predição das taxas de ocupação de leitos de UTI no Brasil e em suas regiões em decorrência da COVID-19 utilizando *Support Vector Regression*. **Métodos:** Inicialmente, foi realizada a predição da taxa de vacinação para utilizá-la como parâmetro para a predição da taxa de ocupação dos leitos de UTI no Brasil e em suas regiões, para 7, 14, 21 e 28 dias. A técnica *Grid Search* foi utilizada para definição dos hiperparâmetros do modelo. **Resultados:** O modelo usando o *kernel rbf* obteve os melhores resultados preditivos, portanto foi o utilizado. O erro obtido, no geral, foi satisfatório, onde as predições apontaram para um declínio na ocupação dos leitos de UTI. **Conclusão:** Os resultados desta pesquisa apontam que deve haver uma significativa redução da ocupação dos leitos de UTI conforme mais pessoas forem completamente imunizadas, tanto no Brasil como um todo quanto em suas regiões.

### ABSTRACT

**Keywords:** ICU beds; *Machine Learning*; COVID-19

**Objectives:** To propose a *Machine Learning* model to predict the ICUs bed occupancy rate Brazil caused by COVID-19 in Brazil and its regions using *Support Vector Regression* technique. **Methods:** Initially, the prediction of vaccination rate was made in order to use its results as a parameter to predict the ICUs bed occupancy rate in Brazil and its regions for 7, 14, 21 and 28 days ahead. The *Grid Search* technique was used to determine the hiperparameters of the model. **Results:** The model using the *rbf kernel* obtained the best results, thus it was selected. The general error was satisfactory, and the prediction pointed to a decline in ICUs bed occupancy rate. **Conclusion:** The results of this research indicate that there should be a significant reduction in ICUs bed occupancy as more people are fully immunized, both in Brazil as a whole and in its regions.

### RESUMEN

**Descriptores:** Camas de UCI; *Machine Learning*; COVID-19

**Objetivos:** Proponer un modelo de *Machine Learning* para predecir las tasas de ocupación de camas de UCI en Brasil y sus regiones debido al COVID-19 utilizando la *Regresión de Vectores de Apoyo*. **Métodos:** Inicialmente, se predijo la tasa de vacunación para utilizarla como parámetro para predecir la tasa de ocupación de camas de UCI en Brasil y sus regiones durante 7, 14, 21 y 28 días. Para definir los hiperparámetros del modelo se utilizó la técnica de búsqueda en cuadrícula. **Resultados:** El modelo que utiliza el *kernel rbf* obtuvo los mejores resultados de predicción, por lo que fue el utilizado. El error obtenido, en general, fue satisfactorio, ya que las predicciones apuntaban a un descenso en la ocupación de las camas de la UCI. **Conclusión:** Los resultados de esta investigación indican que debería haber una reducción significativa de la ocupación de camas en las UCI a medida que más personas se vacunan completamente, tanto en Brasil en su conjunto como en sus regiones.

<sup>1</sup> Bacharel em Ciência e Tecnologia, Universidade Federal Rural do Semi-Árido, Pau dos Ferros (RN), Brasil

<sup>2</sup> Estudante de Ciência e Tecnologia, Universidade Federal Rural do Semi-Árido, Pau dos Ferros (RN), Brasil.

<sup>3</sup> Professor do Departamento de Engenharias e Tecnologia da Universidade Federal Rural do Semi-Árido, Pau dos Ferros (RN), Brasil.

## INTRODUÇÃO

De acordo com a Organização Mundial da Saúde<sup>(1)</sup>, em dezembro de 2019, foram reportados na cidade de Wuhan, na China, os primeiros casos de uma pneumonia de causas até então desconhecidas, onde em 3 de janeiro de 2020 já foram registrados 44 casos acumulados da doença. Posteriormente, o novo coronavírus SARS-CoV-2 veio a ser apontado como o agente responsável pelo surto, e, em 11 de março de 2020, veio a ser declarado como pandemia pela OMS, após vários novos casos serem reportados em diversos países do mundo<sup>(2)</sup>.

Os enormes desafios trazidos pela pandemia do coronavírus têm afetado as mais diversas áreas de todos os países do mundo, indo desde a educação, economia em geral e, em especial, a área da saúde, onde os desafios são implacáveis e se ampliam a todas as áreas do cuidado. Nesse contexto, é de fundamental importância para os serviços de saúde conseguir antecipar os riscos e avanços da doença para que assim seja possível gerenciar as ações preventivas na assistência e cuidado<sup>(3)</sup>.

No final de julho de 2021, o número cumulativo de mortes pela doença já ultrapassava os 4 milhões, enquanto o número de casos acumulados de COVID-19 no mundo já era superior a 190 milhões<sup>(4)</sup>.

No Brasil, desde o primeiro caso confirmado de COVID-19 até o início de julho de 2021, o país havia acumulado mais de 18,5 milhões de casos confirmados de COVID-19 e cerca de 520 milhões de mortes pela doença<sup>(5)</sup>, sendo o país mais afetado pela pandemia na América Latina e figurando entre os três países mais afetados no mundo, sendo o país com o segundo maior número de mortes, atrás apenas dos Estados Unidos, e o terceiro país em relação ao número de casos confirmados da doença, ficando atrás dos Estados Unidos e Índia<sup>(4)</sup>.

Como consequência do elevado número de casos no país, há uma elevada taxa de ocupação dos leitos de Unidade de Terapia Intensiva (UTI), onde, no início de maio de 2021, das 27 unidades federativas existentes no Brasil, 19 estavam com mais de 60% de sua ocupação<sup>(6)</sup>.

Diante das atuais dificuldades enfrentadas no Brasil, tendo grande parte de seus leitos de UTI ocupados, e tendo em vista a capacidade e importância de predição das técnicas de *Machine Learning*, o presente trabalho tem como objetivo realizar e analisar a predição das taxas de ocupação futuras de leitos de UTI no Brasil em decorrência da COVID-19, utilizando como técnica preditiva o *Support Vector Regression* (SVR) e, para auxiliar no ajuste dos hiperparâmetros, foi usado o algoritmo Grid Search. Como métricas para avaliação do modelo foram usados o erro médio absoluto (MAE) e erro quadrático médio (RMSE).

O presente artigo está organizado da seguinte maneira: a Seção 2 apresenta os Trabalhos Relacionados ao presente estudo, enquanto a Seção 3 trata dos Materiais e Métodos usados no desenvolvimento do trabalho. A Seção 4 traz os resultados e discussões e, por fim, a Seção 5 apresenta as conclusões.

## TRABALHOS RELACIONADOS

A presente seção irá tratar dos principais trabalhos relacionados ao uso da técnica SVR para predição de eventos relacionados a COVID-19.

O trabalho de Shahid, Zameer e Muneeb<sup>(7)</sup> teve como objetivo comparar o desempenho de técnicas de previsão temporal, na predição dos casos confirmados, mortes e recuperados da COVID-19 em dez países do mundo. As técnicas comparadas foram a *Autoregressive Integrated Moving Average* (ARIMA), *Support Vector Regression* (SVR) com *kernels* polinomiais e gaussiano, e as redes neurais recorrentes *Long Short-term Memory* (LSTM), *Bi-directional Long Short Term Memory* (Bi-LSTM), *Gated Recurrent Unit* (GRU). Se tratando do Brasil, as técnicas ARIMA e SVR não performaram bem para os dados utilizados, com alto erro de predição nas métricas MAE e RMSE e valores negativos do  $R^2$ , já as técnicas LSTM, Bi-LSTM e GRU apresentaram bons resultados no geral.

O trabalho de Hazarika e Gupta<sup>(8)</sup> propõe um modelo de redes de enlaces funcionais de vetor aleatório (RVFL) acoplado a *Wavelet* (WCRVFL). Os dados decompostos da *wavelet* são fornecidos diretamente como uma entrada para o RVFL e os dados dos 5 países mais afetados pela pandemia até julho de 2020 são fornecidos como uma entrada para o modelo. Uma previsão dos casos totais de infectados pela COVID-19 de 60 dias à frente é retratada para cada país usando RVFL e WCRVFL e comparadas com a SVR, onde o WCRVFL obtém melhores resultados nas métricas RMSE e  $R^2$ , usando a função Sigmoid (0.00055 e 0.99996, respectivamente). A SVR, no entanto, também apresentou um bom  $R^2$  (0.99958), porém um RMSE ligeiramente maior que a WCRVFL (0.02386).

O trabalho de da Silva et al.<sup>(9)</sup> buscou fazer a comparação entre 5 modelos convencionais e 5 híbridos com o *Variational Mode Decomposition* (VMD), que foram empregados para a predição de casos da COVID-19 no Brasil e nos EUA usando variáveis climáticas, onde as variáveis de temperatura e precipitação são usadas como atributos na predição. Os resultados a respeito do Brasil estão divididos entre 5 estados: Amazonas (AM), Ceará (CE), Rio de Janeiro (RJ), São Paulo (SP) e Pernambuco (PE). No AM, a técnica VMD-BRNN obteve os melhores resultados, já no CE, RJ e SP, VMD-Cubist alcançou as melhores predições, por fim, em PE o modelo Cubist e a SVR apresentaram as melhores performances.

O trabalho de Ribeiro et al.<sup>(10)</sup> buscou explorar e comparar a capacidade preditiva de técnicas de regressão usando *Machine Learning* e modelos estatísticos para predizer os casos acumulados da COVID-19 no Brasil em até seis dias à frente, onde foram usados como entrada dos modelos dados de dez estados com alta incidência de casos confirmados. Como resultado, SVR e *Stacking Ensemble* apresentam os melhores resultados preditivos para a maioria dos estados.

Nos trabalhos citados, o SVR foi usado na sua modelagem convencional ou híbrido com outras técnicas, com objetivos distintos de predição, sendo que em apenas um dos trabalhos<sup>(11)</sup> houve um estudo focado apenas na

técnica SVR. Na maioria dos estudos, o SVR performou satisfatoriamente bem para prever dados relacionados a COVID-19, mas vale salientar que nenhum dos trabalhos abordados tratou da predição da taxa de ocupação de leitos de UTI, seja dentro ou fora do Brasil.

## MATERIAIS E MÉTODOS

### Dataset

Para o desenvolvimento desta pesquisa, foi necessária a utilização de dados relacionados tanto à taxa de ocupação dos leitos de UTI, quanto à taxa de vacinação no Brasil. Desse modo, distintos bancos de dados foram utilizados, os quais podem ser encontrados em FIOCRUZ (2021).

O primeiro *dataset* trata de uma série temporal da evolução da taxa de ocupação de leitos de UTI da COVID-19 em todas as unidades federativas do Brasil. Na data em que foi realizado o download do *dataset*, o mesmo possuía 37 instâncias, com início em 17 de julho de 2020 e término em 19 de julho de 2021. Os atributos do banco de dados são: a data da coleta do dado (*data\_coleta*), a Unidade Federativa (*uf*), a taxa de ocupação dos leitos de UTI (*taxa\_ocup*) e o nível de alerta da taxa de ocupação (*alerta*).

Os demais *datasets* trazem uma série temporal a respeito do percentual de pessoas vacinadas contra a COVID-19 com a segunda dose ou dose única, e consistem em bancos de dados específicos para cada unidade federativa do Brasil ou para o país propriamente dito. Todos os dados de vacinação referentes às UFs foram coletados no intervalo de tempo de 17 de janeiro de 2021 a 19 de julho de 2021. Os *datasets* obtidos apresentam os mesmos atributos, sendo eles: a data de coleta dos dados (*DataBrasil*), o percentual de vacinação no Brasil (*PercentualBrasil*), a data de coleta por UF (*Data<nome\_da\_UF>*) e o percentual de vacinação por UF (*Percentual<nome\_da\_UF>*).

Na fase de tratamento de dados, como este trabalho busca prever a taxa de ocupação dos leitos de UTI no Brasil e em cada uma de suas cinco regiões, fez-se necessário agrupar os dados de cada UF de acordo com sua região, em seguida, tanto o percentual da taxa de vacinação quanto da taxa de ocupação de leitos de UTI em cada região foi calculado a partir da média aritmética das taxas das respectivas UFs, o que apresenta uma estimativa satisfatória do valor real. É importante frisar que, nos bancos referentes às taxas de vacinação, a coleta dos dados foi feita diariamente, logo, a fim de manter todos os dados em conformidade, foram levados em consideração os dados presentes nos mesmos intervalos de tempo entre as coletas disponíveis no *dataset* da taxa de ocupação dos leitos.

Dentre os atributos do *dataset* referente a taxa de ocupação dos leitos de UTI, após o tratamento inicial, apenas a “*data\_coleta*” e a “*taxa\_ocup*” foram utilizados, sendo o atributo UF substituído por “*regiao*”, já nos bancos de dados referentes às taxas de vacinação, o percentual, foi adicionado um novo atributo “*regiao*”, ao passo que o “*Percentual<nome\_UF>*” passou a ser

apenas “*taxa\_vacinacao*”.

Para realizar a predição da taxa de ocupação dos leitos levando em consideração a taxa de vacinação, esta última deve se tornar um atributo, assim, um novo banco de dados foi criado a partir da junção dos *datasets* citados anteriormente, onde os atributos utilizados foram a data da coleta (*data\_coleta*), a região ou o Brasil como um todo (*regiao*), o percentual de pessoas totalmente vacinadas (*taxa\_vacinacao*) e a taxa de ocupação de leitos de UTI pela COVID-19 (*taxa\_ocup*).

Ao final do processamento, o novo banco de dados foi dividido em seis novos *datasets* contendo 37 instâncias cada, sendo um banco para cada região do país e um com os dados referentes ao Brasil como um todo.

### Implementação em Python

#### Bibliotecas utilizadas

Para realizar tanto a implementação da SVR quanto a importação dos *datasets*, foram utilizadas as bibliotecas da linguagem Python Pandas, Numpy, Matplotlib, Plotly e Sklearn. A biblioteca Sklearn possui uma classe chamada SVR, acessada através do comando `sklearn.svm.SVR()`, que permite a implementação da técnica de *Machine Learning Support Vector Regression*. Além disso, possui também a classe `GridSearchCV`, acessível a partir de `sklearn.model_selection.GridSearchCV()`, que traz a implementação da técnica *Grid Search*, cuja função é a otimização de hiperparâmetros.

#### Ajuste dos hiperparâmetros

A classe SVR, da biblioteca Sklearn, possui diversos parâmetros que podem receber inúmeros valores distintos, dentre eles o *kernel*, onde, nesta pesquisa, dois foram levados em consideração: a “*rbf*” (função de base radial) e o “*poly*” (polinomial), por serem os que podem apresentar variações de comportamento mais generalista.

Com a finalidade de ajustar os hiperparâmetros de forma automatizada, a técnica *Grid Search* foi utilizada. No entanto, quando o número de possibilidades para um parâmetro é muito alto, a *Grid Search* gera um grande esforço computacional, por esse motivo, os testes para os *kernels* “*poly*” e “*rbf*” foram feitos separadamente. No caso do *kernel* “*poly*”, o ajuste do hiperparâmetro “*degree*” (grau) também foi feito isoladamente para os valores 5, 6 e 8. Como a complexidade da função é proporcional ao grau do polinômio, os valores foram escolhidos buscando evitar funções muito simples (*underfitting*) e também funções muito complexas (*overfitting*). Inicialmente, foram realizados testes para os bancos de dados da percentagem de vacinação e, posteriormente, os testes foram aplicados aos dados de percentual de ocupação de leitos de UTI.

#### Metodologia da predição

Para realizar a predição da taxa de ocupação dos leitos de UTI pela COVID-19 para 7, 14, 21 e 28 dias no futuro, a taxa de vacinação nessas datas precisa ser conhecida. No entanto, como são datas futuras, foi necessário realizar uma predição do percentual de vacinação para tais datas, antes de prever as taxas de ocupação dos leitos. Desse

modo, a SVR foi implementada primeiramente para prever os dados de vacinação e, a partir desses resultados, a predição da taxa de ocupação de leitos foi realizada.

## RESULTADOS E DISCUSSÕES

Esta seção detalha os resultados da seleção dos hiperparâmetros da SVR e das predições realizadas tanto para o Brasil quanto para cada região do país separadamente.

### Hiperparâmetros selecionados

Para realizar a predição da taxa de ocupação de leitos de UTI no Brasil e em suas regiões devido a COVID-19, além das datas presentes nos bancos de dados, foi adicionada a taxa de vacinação nas mesmas datas como atributo. No entanto, como a taxa de vacinação estava disponível apenas até o dia 22 de julho de 2021, foi necessário realizar a sua predição para os mesmos períodos pretendidos para a taxa de ocupação dos leitos (7, 14, 21 e 28 dias), para que, então, esses valores fossem usados como atributos de entrada juntamente com as datas.

O algoritmo de *Grid Search* foi aplicado inicialmente apenas para o banco de dados contendo a série temporal da taxa de vacinação no Brasil, sendo executados 5 testes variando o grau (no caso do *kernel* polinomial), o *kernel* e o valor do parâmetro C, este último definido de forma empírica, cujos resultados estão dispostos na Tabela 1.

A métrica *Mean Squared Error* (MSE) foi utilizada para definição do *score*, desse modo, quanto mais o valor se aproxima de zero, melhor, ou seja, o teste 4 apresentou o melhor resultado.

### Predição da taxa de vacinação

Com os valores dos hiperparâmetros definidos de

acordo com o resultado do teste 4 apresentado na subseção 4.1, as taxas de vacinação para 7, 14, 21 e 28 dias a frente foram previstas para que servissem como atributos da taxa de ocupação de leitos de UTI. O processo de predição foi realizado tanto para os dados referentes ao Brasil quanto para cada região separadamente, com a finalidade de obter os atributos necessários para a predição da taxa de ocupação dos leitos de UTI. As figuras e tabelas referentes a predição da taxa de vacinação no Brasil e em cada uma de suas regiões encontram-se em um repositório online\*.

### Predição para o Brasil

A partir dos valores preditos para a taxa de vacinação, foi realizada a predição da taxa de ocupação dos leitos nas UTIs do país, pela COVID-19. A Figura 1 apresenta o gráfico da taxa de ocupação em relação aos dias.

Para 7 dias no futuro, a predição da taxa de ocupação dos leitos de UTI no país foi de 54.91%, para 14 dias a taxa foi de 49.14%, caindo para 42.86% 21 dias no futuro e para 36.12% ao completar 28 dias. O valor obtido da MAE foi de 3.4081, e a RMSE foi de 4.3825.

Os resultados apontam para uma queda na taxa de ocupação de leitos de UTI no país conforme o tempo passa e a taxa de vacinação aumenta, tendo uma diminuição de 18.79% após 28 dias. Os valores da MAE e RMSE foram significativamente maiores quando comparados com a predição da taxa de vacinação isolada, o que pode ter ocorrido devido ao fato de haver dois atributos (a taxa de vacinação e as datas de coleta) ao invés de apenas um (data de coleta).

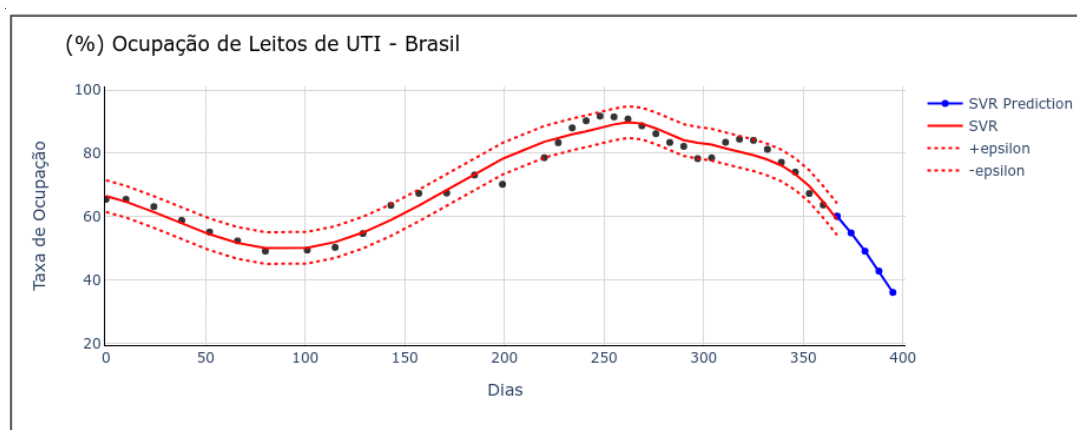
### Predição para a região Norte

A região norte apresenta uma taxa de vacinação real

**Tabela 1** – Testes do Grid Search para taxa de vacinação no Brasil

Teste	kernel	C	degree	coef0	gamma	score
1	poly	1	5	1	auto	0.12097
2	poly	100	6	1	auto	-0.14301
3	poly	100	8	1	auto	0.10162
4	rbf	10000	-	0	auto	0.06771
5	rbf	10000	-	0	scale	0.28718

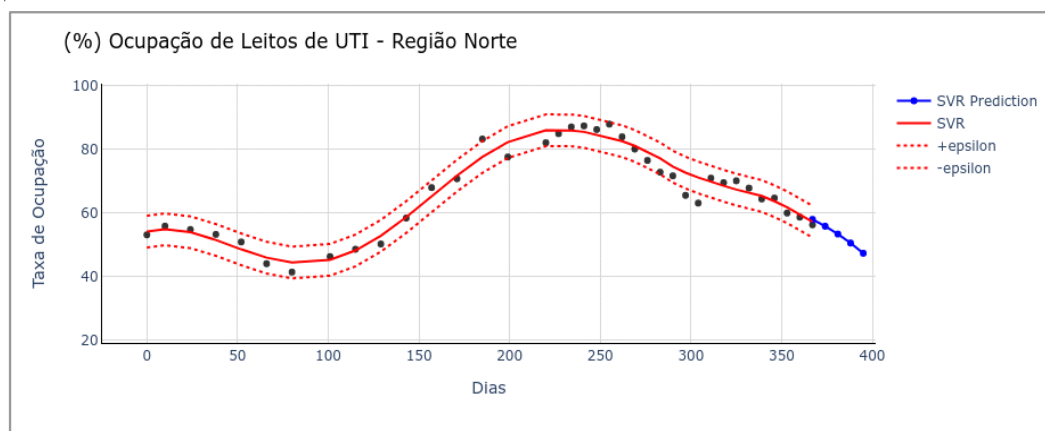
Fonte: Os autores



Fonte: Os autores

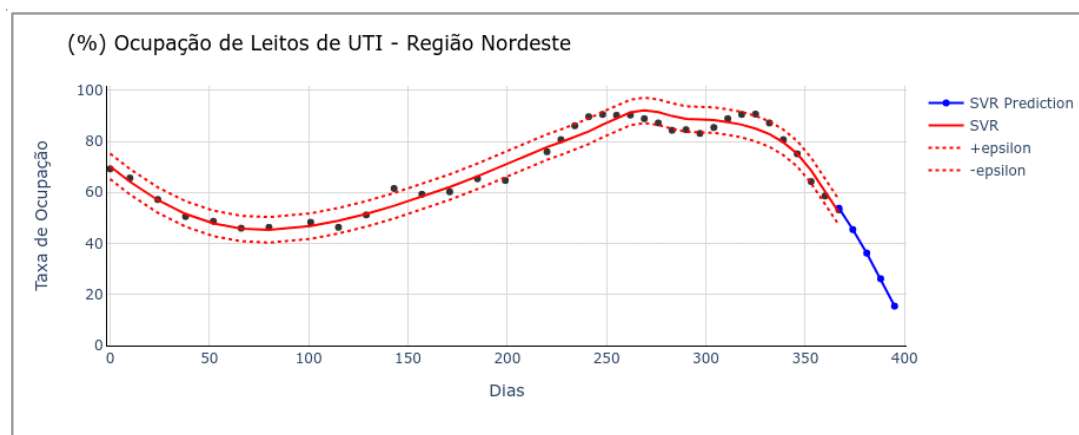
**Figura 1** - Predição da taxa de ocupação de leitos de UTI no Brasil

\* Repositório: <https://github.com/gabrielgcb/ICU-beds-COVID19-Brazil>



Fonte: Os autores

**Figura 2** - Predição da taxa de ocupação de leitos de UTI no Norte



Fonte: Os autores

**Figura 3** - Predição da taxa de ocupação de leitos de UTI no Nordeste

baixa se comparada a do Brasil, sendo de apenas 18.36% após 28 dias, ou seja, apresenta uma tendência de aumento de apenas 0.74% 28 dias no futuro, com um MAE de 0.8764 e uma RMSE de 0.9459. A Figura 2 dispõe da taxa de ocupação de leitos de UTI por COVID-19 na região Norte em relação ao tempo, levando em consideração também a taxa de vacinação.

A predição para a região norte gerou os seguintes resultados: 7 dias: 55.84%; 14 dias: 53.38%; 21 dias: 50.55% e para 28 dias: 47.32%. A MAE foi de 4.2266 e a RMSE foi de 5.3936, representando valores um pouco maiores que na predição feita para o Brasil. A predição mostra que, para 28 dias no futuro, é esperada uma queda de 8.52% na taxa de ocupação dos leitos de UTI na região Norte do país em relação a última data presente no banco de dados.

#### Predição para a região Nordeste

A taxa de vacinação da região nordeste é próxima a do Brasil, e, para os dias futuros, os resultados mostram um aumento de 2.41%. Tanto o MAE (0.8097) quanto a RMSE (0.9650) para a predição no Nordeste, foram significativamente satisfatórios, pois apresentaram um valor baixo.

A partir da taxa de vacinação real e predita, a taxa de ocupação de leitos de UTI para 7, 14, 21 e 28 dias foi encontrada, apresentando uma redução de mais de 30% no período citado, sendo: 7 dias: 45.39%; 14 dias: 36.14%; 21 dias: 26.09% e para 28 dias: 15.36%.

O valor do MAE e da RMSE foi mais baixo que o da região Norte e que o do Brasil, sendo 2.7322 e 3.2081, respectivamente, ou seja, o erro médio da predição para a região Nordeste foi menor que o das citadas. A Figura 3 ilustra os resultados da região Nordeste.

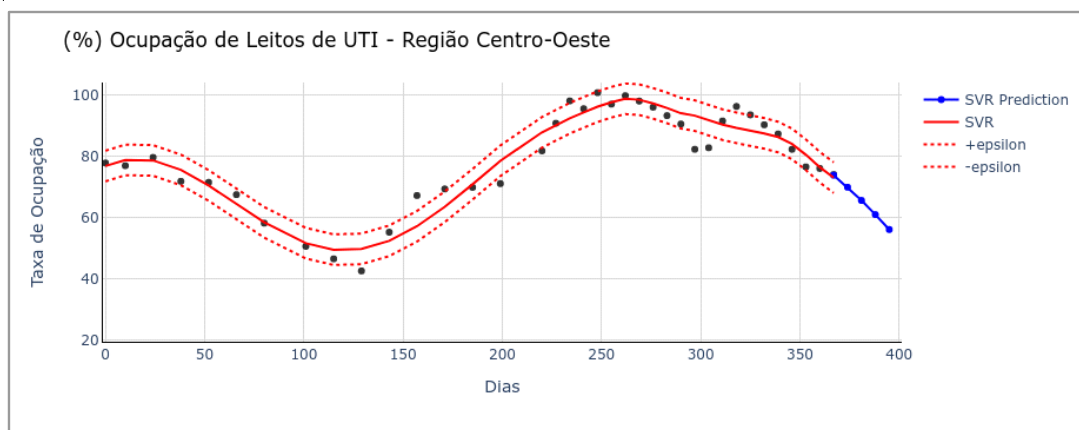
#### Predição para a região Centro-Oeste

Dentre as cinco regiões do Brasil, a Centro-Oeste apresenta uma das maiores taxas de imunização total da população, dado que se reflete na taxa prevista para os dias futuros, apresentando um crescimento de 9.74% 28 dias no futuro, com um MAE de 0.9801 e uma RMSE de 1.1901.

Com a expectativa de um avanço significativo na taxa de vacinação na região, a taxa de ocupação de leitos de UTI prevista apresenta uma redução de 13.81%, como ilustra o gráfico disponível na Figura 4, onde os resultados foram: para 7 dias: 69.86%; 14 dias: 65.55%; 21 dias: 60.93% e 28 dias: 56.05%. O MAE obtido foi de 3.7979 e a RMSE, 4.4497.

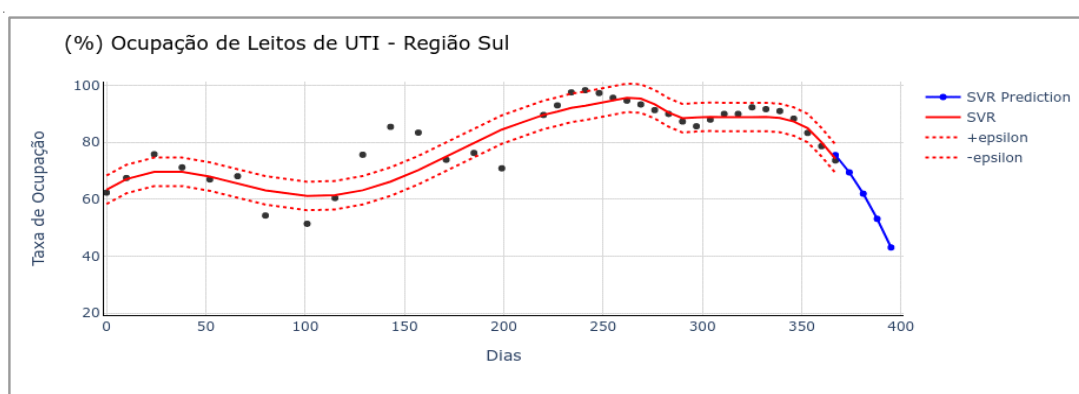
#### Predição para a região Sul

A região Sul apresenta a maior expectativa de aumento na taxa de imunização total de indivíduos no Brasil, com 11.61% para 28 dias no futuro, chegando a 36.61% da população totalmente vacinada. O valor do MAE (0.8019) e da RMSE (0.9565) foi significativamente baixo, isto é, a predição da taxa de vacinação na região Sul apresentou um erro médio baixo.



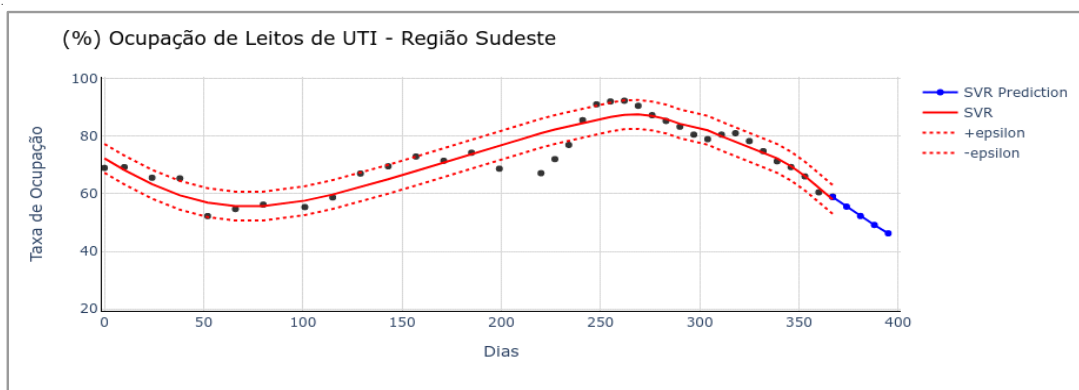
Fonte: Os autores

**Figura 4** - Predição da taxa de ocupação de leitos de UTI no Centro-Oeste



Fonte: Os autores

**Figura 5** - Predição da taxa de ocupação de leitos de UTI no Sul



Fonte: Os autores

**Figura 6** - Predição da taxa de ocupação de leitos de UTI no Sudeste

O avanço na vacinação reflete na taxa de ocupação de leitos de UTI na região pela COVID-19, cujo resultado apresenta uma redução de 21.36% no intervalo de tempo definido. O gráfico disposto na Figura 5 apresenta a evolução da taxa de ocupação dos leitos em relação ao tempo, cujo resultado para 7 dias foi 69.47%; para 14 dias, 62.00%; 21 dias, 53.17% e para 28 dias, 43.11%

O MAE obtido foi 7.9167, enquanto a RMSE foi de 11.0068. O erro foi significativamente alto e maior que o da predição realizada para dados do Brasil e das demais regiões do país, o que implica dizer que o erro médio da predição da taxa de ocupação de leitos de UTI para a região Sul apresentou um valor alto se comparado ao das predições para as outras regiões, ao contrário da

predição da taxa de vacinação.

#### Predição para a região Sudeste

Os resultados da predição para a região Sudeste sugerem um aumento de 2.72% da taxa de vacinação nos 28 dias à frente, com um erro médio baixo, onde o MAE = 0.7834 e a RMSE = 0.9804.

Mesmo com o aumento relativamente pequeno na taxa de vacinação, o modelo prevê uma redução de mais de 9% na ocupação de leitos de UTI na região no período definido, apresentando, também, o MAE e a RMSE mais baixos dentre as regiões e o país como um todo, sendo 4.7375 e 6.1222, respectivamente. Os resultados da predição da taxa de ocupação foram: 7 dias: 55.53%; 14

dias: 52.28%; 21 dias: 49.17%; 28 dias: 46.27%.

A Figura 6 dispõe de um gráfico que relaciona a taxa de ocupação de leitos de UTI pela COVID-19 no Sudeste em relação ao tempo, levando também em consideração a taxa de vacinação real e predita como atributo da predição.

## CONCLUSÃO

A pandemia do novo coronavírus, que teve seu início em Wuhan, na China, causou milhares de mortes em todo o mundo desde seu início, que foi no ano de 2020. No Brasil, no final de julho de 2021, havia quase 20 milhões de casos acumulados de COVID-19 e mais de 550 mil mortes pela doença. Devido ao impacto da doença, vários pesquisadores propuseram modelos de *Machine Learning* para prever dados a respeito do coronavírus no Brasil e no mundo.

O grande número de casos de COVID-19 levou a uma alta taxa de ocupação dos leitos de UTI no Brasil, o que pode gerar um caos na saúde pública. No entanto,

no ano de 2021 iniciou-se a vacinação no país, fato este que busca imunizar o máximo possível de pessoas contra a COVID-19, reduzindo a propagação da doença e, conseqüentemente, a taxa de ocupação de leitos de UTI. Desse modo, com o objetivo de prever o impacto da vacinação na ocupação dos leitos, este trabalho implementou a técnica de *Machine Learning* SVR, com dados referentes a taxa de vacinação e as datas de coleta, para prever a taxa de ocupação dos leitos de UTI pela COVID-19 no Brasil e em cada uma de suas regiões para 7, 14, 21 e 28 dias no futuro.

Os resultados desta pesquisa mostram que deve haver uma redução significativa da ocupação dos leitos de UTI conforme mais pessoas forem completamente imunizadas, tanto no Brasil como em todo quanto em suas regiões de forma mais específica, com destaque positivo para região Nordeste, onde é prevista uma redução de mais de 30% na ocupação dos leitos. Esses resultados reforçam a importância da vacinação e da imunização da população no enfrentamento da pandemia da COVID-19, e podem auxiliar os governos no combate ao coronavírus.

## REFERÊNCIAS

1. OMS. Pneumonia of unknown cause – China. 2019. Disponível em: <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>. Acesso em 14 mai. 2021.
2. Universidade Aberta do Sistema Único de Saúde, UNASUS. Disponível em: <https://www.unasus.gov.br/noticia/organizacao-mundial-de-saude-declara-pandemia-de-coronavirus>. Acesso em: 23 set. 2021.
3. Espinosa M. M.; Oliveira E. C. de; Melo J. S.; Damasceno R. D.; Terças-Trettel A. C. P. Predição de casos e óbitos de COVID-19 em Mato Grosso e no Brasil. *Journal of Health & Biological Sciences* [internet]. Mai 2020 [citado em 06 out 2020];8(1):[7 p.]. Disponível em: <https://periodicos.unichristus.edu.br/jhbs/article/view/3224/1111>.
4. OMS. WHO Coronavirus (COVID-19) Dashboard. 2021. Disponível em: <https://covid19.who.int>. Acesso em: 24 set. 2021.
5. Ritchie at. Al. Our World in Data. Coronavirus Pandemic (COVID-19). Disponível em: <https://ourworldindata.org/coronavirus>. Acesso em: 24 set. 2021.
6. FIOCRUZ. Instituto de Comunicação e Informação Científica e Tecnológica em Saúde (ICICT). MonitoraCovid-19. Rio de Janeiro, 2020. Disponível em: <https://bigdata-covid19.icict.fiocruz.br/>. Acesso em: 24 set. 2021.
7. Shahid F, Zameer A, Muneeb M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals*. Aug 19, doi: <https://doi.org/10.1016/j.chaos.2020.110212>.
8. Hazarika BB, Gupta D. Modelling and forecasting of COVID-19 spread using wavelet-coupled random vector functional link networks. *Applied Soft Computing*. 2020 Ago 13, doi: <https://doi.org/10.1016/j.asoc.2020.106626>.
9. da Silva RG et al. Forecasting Brazilian and American COVID-19 cases based on artificial intelligence coupled with climatic exogenous variables. *Chaos, Solitons & Fractals*. 2020 Jun 30, doi: <https://doi.org/10.1016/j.chaos.2020.110027>.
10. Ribeiro MHD et al. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals*. 2020 Mai. 1, doi: <https://doi.org/10.1016/j.chaos.2020.109853>.
11. Penga Y, Nagata MH. An empirical overview of nonlinearity and overfitting in machine learning using COVID-19 data. *Chaos, Solitons and Fractals*. 2020 Jun 30. <https://doi.org/10.1016/j.chaos.2020.110055>.