

Aplicação de Redes de Deep Learning Recurrent Neural Network, Long Short-Term Memory e Gated Recurrent Unit na Predição da COVID-19 no Cenário Pós-Vacinação

Rafaella Silva Ferreira¹, Wallace Casaca¹, Marilaine Colnago¹

¹Instituto de Biociências, Letras e Ciências Exatas
Departamento de Ciências de Computação e Estatística
Universidade Estadual Paulista (UNESP) – São José do Rio Preto, SP – Brasil

{rafaella.ferreira, wallace.casaca, marilaine.colnago}@unesp.br

Abstract. *This work proposes the adaptation of three artificial neural networks to predict COVID-19 time series in Brazil, considering the current scenario of data scarcity and recurrent waves of the disease, though of lesser magnitude compared to 2020 and 2021. The main goal is to assess the performance of different neural network architectures in forecasting daily COVID-19 cases in the state of São Paulo. The evaluated architectures are as follows: Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Gated Recurrent Unit (GRU). The performance of each model was analyzed in terms of adherence to real data and the ability to capture complex temporal patterns, in a context of sudden increases and declines in cases of lesser severity due to vaccination. The results demonstrated high accuracy of the trained networks and provided insights to improve the quality of the predictions, which are essential for pandemic control strategies, especially during periods of disease resurgence. Thus, this work aims to contribute to the advancement of predictive neural network applications for COVID-19, particularly in the post-vaccination context.*

Resumo. *Este trabalho propõe a adaptação de três redes neurais artificiais para prever séries temporais da COVID-19 no Brasil, considerando o atual cenário de escassez de dados e ondas recorrentes da doença, mas de menor magnitude comparado a 2020 e 2021. O principal objetivo é avaliar o desempenho de diferentes arquiteturas de redes neurais na previsão dos casos diários de COVID-19 no estado de São Paulo. As arquiteturas exploradas foram: Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM) e Gated Recurrent Unit (GRU). O desempenho de cada modelo foi analisado quanto à aderência aos dados reais e à capacidade de capturar padrões temporais complexos, em um contexto de aumentos-e-declínios súbitos de casos de menor gravidade devido à vacinação. Os resultados demonstraram alta precisão das redes treinadas e forneceram diferentes perspectivas para melhorar a qualidade das previsões, essenciais para estratégias de controle da pandemia, especialmente durante os períodos de recrudescimento da doença. Assim, este trabalho visa contribuir para o avanço de aplicações de redes neurais preditoras da COVID-19, particularmente no contexto pós-vacinação.*

1. Introdução

Em dezembro de 2019, a humanidade testemunhou o surgimento do coronavírus SARS-CoV-2, em Wuhan, China. A doença, denominada COVID-19, adquiriu status de surto no início de 2020, e rapidamente evoluiu para uma pandemia em março de 2020, resultando em mais de 7 milhões de óbitos até então [Roosa et al. 2020, Worobey 2021, Worldometer 2024]. A pandemia não só resultou em uma crise sanitária global, mas também transformou a economia e o cotidiano das pessoas, levando a emergências sanitárias devido à sobrecarga nos sistemas hospitalares, especialmente em 2020 e 2021.

Apesar da importância da monitoração contínua dos índices epidemiológicos, tal prática tem enfrentado obstáculos no Brasil, o que inclui o desinteresse do poder público em manter mobilizada suas equipes técnicas nas tarefas de coleta e atualização dos dados na fase atual da doença. A ausência de políticas públicas e de monitoramento sistemático da situação pandêmica atual, bem como escassez de atualizações nos dados tem inviabilizado análises do tipo *data-driven*, que anteriormente eram capazes de capturar a realidade da pandemia [Oliveira 2024]. Além disso, a inexistência de campanhas de incentivo à testagem e a propagação de *fake news* contribuem para a subnotificação nos dados, tornando a avaliação da doença um desafio para a saúde pública nacional, considerando o efeito sanfona da COVID-19 observado em 2022 e 2023 [Mendes 2024, Capomaccio 2024].

A capacidade de prever a disseminação e comportamento futuro do vírus é crucial para a implementação de políticas públicas efetivas e intervenções de natureza preventivas. Nesse contexto, as redes neurais tem se estabelecido como uma ferramenta poderosa na análise de séries temporais e predição de casos e óbitos da COVID-19, uma vez que tais arquiteturas são capazes de capturar com precisão padrões complexos e não lineares nos dados, estando habilitadas para lidar com grandes volumes de dados e adaptando-se com exatidão à novas instâncias após a fase de treinamento.

Considerando o contexto e desafios destacados acima, o presente trabalho tem como objetivo avaliar o desempenho de diferentes arquiteturas de redes neurais na predição dos casos diários de COVID-19 no estado de São Paulo. Tal investigação é fundamental não apenas para compreender o desempenho dos modelos na captura das variações nos dados epidemiológicos, mas também para pavimentar a via para pesquisas futuras, sobretudo no contexto brasileiro pós-vacinação, onde a ausência de dados completos e acurados constitui um desafio permanente. Em particular, foram implementadas e exploradas as seguintes redes: *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), e *Gated Recurrent Unit* (GRU). Além disso, este trabalho contribui para o monitoramento sistemático dos casos da COVID-19 em um contexto pós-pandemia, tema que não tem sido amplamente coberto na literatura. A pesquisa também avança na aplicação e adaptação de uma rede do tipo GRU para prever casos de COVID-19, cuja arquitetura foi ainda pouco explorada na literatura especializada sobre a doença.

1.1. Trabalhos Relacionados

Esta seção aborda a importância da utilização de redes neurais no enfrentamento da COVID-19, discutindo trabalhos com o mesmo propósito, principalmente em obras focadas na aplicação de redes recorrentes para prever novos casos da doença.

Em [Hawas 2020], os autores conduziram um estudo com dados do Brasil, empregando um modelo baseado em RNN combinado com autoencoders do tipo GRU-CNN.

No referido trabalho, o modelo aplicado apresentou resultados satisfatórios em janelas de curto prazo, entretanto, encontrou dificuldades na construção de previsões a longo prazo (horizonte mensal). Os autores também enfatizaram a importância do *tuning* dos modelos analisados, e a necessidade de considerar o contexto específico de cada país. Já em [Zeroual et al. 2020], foi realizado um estudo comparativo envolvendo cinco métodos de aprendizado profundo: RNN, LSTM, GRU, BiLSTM (*Bidirectional Long Short-Term Memory*) e VAE (*Variational Autoencoder*). Tais métodos foram aplicados na previsão global de casos da COVID-19 a partir de um conjunto de dados reduzido, abrangendo seis países: Itália, Espanha, França, China, Estados Unidos e Austrália. Os resultados ressaltaram o potencial promissor dos modelos de aprendizado profundo na previsão dos casos da doença, com o VAE alcançando desempenho superior em relação aos demais.

Na mesma linha dos trabalhos discutidos acima, em [Davahli et al. 2021], os autores aplicaram algoritmos do tipo LSTM e GNN (*Graph Neural Network*), bem como propostas específicas como o GTNN (*Graph-Theory-based Neural Network*) e NGNN (*Neighborhood-based Neural Network*), onde foram utilizadas séries temporais da propagação da doença nos EUA entre 22 de janeiro a 26 de novembro de 2020. Em comparação com o modelo LSTM, o GTNN apresentou melhor desempenho na inferência da disseminação da doença, oferecendo as tendências de aumento e declínio durante uma janela do período investigado. De forma similar, em [Gomez-Cravioto et al. 2021], os autores compararam as redes LSTM e VAR (*Vector Autoregression*) para prever os casos diários da COVID-19 somados de 173 países. Neste caso, o modelo LSTM foi aquele que apresentou melhor acurácia, reduzindo o RMSE (*Root Mean Square Error*) em comparação com o VAR. Ainda sobre uso de redes neurais para a tarefa de previsão das curvas da COVID-19, em [Chandra et al. 2022], foram empregadas diferentes variantes da rede LSTM, incluindo a LSTM bidirecional e LSTM com codificador-decodificador, com foco no avanço da doença em estados da Índia. A partir dos resultados gerados pelos modelos, os autores reportaram existir baixa probabilidade de ocorrência de novas ondas de infecções no último trimestre de 2021, embora destacaram a necessidade de manutenção de vigilância devido ao surgimento de novas variantes do vírus.

Conforme discutido, diferentes propostas tem focado no uso de redes neurais, sobretudo, do tipo LSTM e derivadas, cujo período de análise considerado foram os anos de 2020 e 2021, isto é, os anos de maior descontrole da doença, sem ou baixa cobertura vacinal. Assim, em contraste com as principais obras existentes na literatura, optou-se, neste trabalho, focar na análise de diferentes arquiteturas de redes neurais posto o contexto atual da doença no país, isto é, durante o período de pós-vacinação.

2. Materiais e Métodos

Considerando as discussões apresentadas nas seções anteriores, esta seção introduz as principais fontes de dados e acervos científicos utilizados ao longo da pesquisa, incluindo as etapas de desenvolvimento metodológico e plataformas computacionais utilizadas.

2.1. Fontes de Dados e Acervos Científicos

Com relação aos repositórios científicos empregados, foi realizada a consulta de obras e material bibliográfico nos acervos das editoras IEEE, Springer, Elsevier e MDPI.

Para o treinamento e validação das redes neurais, foram utilizadas séries temporais extraídas de três bases de dados abertas, conforme descrito abaixo:

1. **Worldometer:** Plataforma online, que oferece dados atualizados sobre a pandemia de COVID-19 em todo o mundo, incluindo casos confirmados, óbitos e pessoas recuperadas. Link de acesso em: <https://www.worldometers.info/coronavirus/>.
2. **SEADE:** A *Fundação Sistema Estadual de Análise de Dados*, também conhecida como plataforma SEADE, é uma instituição pública vinculada à Secretaria de Planejamento e Gestão do Estado de São Paulo. Ela disponibiliza dados da COVID-19 em nível municipal, incluindo casos confirmados, óbitos e outras informações relevantes sobre a doença. Link de acesso em: <https://www.seade.gov.br/coronavirus/>.
3. **SIMI-SP:** O *Sistema de Monitoramento Inteligente do Estado de São Paulo*, também conhecido como SIMI-SP, é um acervo de dados para gestão da pandemia. Ela tem tornado público dados sobre internações, ocupações de leitos públicos COVID-19 e outros indicadores. Link de acesso em: <https://www.saopaulo.sp.gov.br/planosp/simi/dados-abertos>.

2.2. Desenvolvimento Metodológico

Visando implementar e validar os modelos de redes neurais explorados neste trabalho, foram realizadas as etapas preliminares de aquisição e tratamento dos dados.

2.2.1. Ferramentas de Análise Exploratória de Dados

A fim de compreender a dinâmica dos dados, foi construído um painel interativo do tipo *Business Intelligence* (BI). Para essa tarefa, foi utilizado o Power BI, plataforma desenvolvida pela Microsoft [Microsoft 2023]. O BI é uma abordagem que permite coletar e analisar dados para apoiar tomadas de decisões, bem como para criar painéis interativos.

A partir do painel construído, foram conduzidas análises com dados globais e, posteriormente, progredindo-se para uma análise mais direcionada aos dados da pandemia no estado de São Paulo. Uma vez que as páginas do painel são totalmente interativas, foi possível uma exploração detalhada da massa de dados. Além disso, a plataforma permitiu imersão total em meio às séries temporais coletadas, proporcionando uma experiência informativa e interativa. O painel interativo de BI criado pode ser acessado pelo link¹.

Além de servir para fins de exploração e compreensão dos dados, o painel construído foi empregado como repositório de dados, concentrando informações relativas à pandemia de março de 2020 até junho de 2023.

2.2.2. Pré-Processamento e Preparação dos Dados

Após análise dos dados, foi necessário a implementação da etapa de preparação dos dados antes da realização dos designs dos modelos de redes neurais.

Inicialmente, os dados coletados das bases já destacadas na subseção 2.1 foram tratados, utilizando-se para tal a linguagem Python, uma das linguagens mais populares na resolução de problemas relacionados à redes neurais [Linge and Langtangen 2020].

¹<http://tiny.cc/covidBI>

Como etapa de pré-processamento dos dados, foram conduzidas as sub-etapas limpeza, transformação e normalização dos dados, similar a [Amaral et al. 2021b]. Além disso, devido à flutuação gerada artificialmente no número de casos de COVID-19 em alguns dias específicos em razão da ausência das atualizações dos dados aos finais de semana e feriados, optou-se pela aplicação da técnica da média móvel (janela de 7 dias), em linha com a metodologia de tratamento de *outliers* adotada por autoridades da área da saúde, como o Ministério da Saúde [Ministério da Saúde 2024].

2.2.3. Redes Neurais Recorrentes

Nesta etapa, foram selecionadas diferentes redes de aprendizado profundo para o tratamento das séries temporais da COVID-19, incluindo os modelos *Recurrent Neural Network* (RNN) [Hewamalage et al. 2021], *Gated Recurrent Unit Neural Network* (GRU) [Chung et al. 2014], e a *Long Short-Term Memory* (LSTM) [Shi et al. 2015]. As três redes foram então exploradas e customizadas para realizar testes com dados abrangendo os anos de 2022 e 2023, com o intuito de comparar o desempenho de cada abordagem diante da escassez de dados e do período de pós-vacinação, onde as ondas da doença passaram a se manifestar de forma menos intensa, porém ainda preocupantes.

A seguir, são apresentadas cada uma das redes utilizadas, incluindo os principais hiperparâmetros ajustados.

1. *Recurrent Neural Network* (RNN)

As RNNs constituem um tipo específico de rede neural profunda projetada para lidar com dados sequenciais, tal como séries temporais. Seu diferencial está na capacidade de considerar não apenas a entrada atual, mas também o contexto das entradas anteriores ao processar cada instância de dados sequencialmente. Isso permite que a rede capture padrões temporais e dependências de longo prazo nos dados.

Para a construção do modelo RNN no contexto da COVID-19, utilizou-se uma camada *SimpleRNN*. O número de neurônios para a referida camada foi calculado através de uma busca ótima de hiperparâmetros (vide subseção 2.2.4), resultando em 184 neurônios para o primeiro experimento, 364 neurônios para o segundo, e 4 para o terceiro. Já a camada de saída do modelo é densa, contendo um neurônio, de forma a condensar a informação aprendida pela camada *SimpleRNN* em uma única previsão, obtendo uma saída definitiva para a próxima amostra da série [Hewamalage et al. 2021].

2. *Long Short-Term Memory* (LSTM)

As redes LSTM correspondem a uma arquitetura do tipo RNN, que comporta determinado grau de liberdade para o comprimento do intervalo, tornando-as altamente adequadas para classificar, processar e prever séries temporais com intervalos de tempo de duração desconhecida.

Para a construção do modelo LSTM a partir dos dados da COVID-19, adotou-se como número de neurônios o valor calculado pela busca ótima de hiperparâmetros, mais precisamente: 152 neurônios para o primeiro experimento, 464 neurônios para o segundo, e 292 para o terceiro. E a camada de saída também foi ajustada como um *layer* denso, com um único neurônio na saída. Para detalhes implementacionais, vide [Shi et al. 2015].

3. Gated Recurrent Unit (GRU)

A arquitetura GRU é outra variante do tipo RNN, a qual foi proposta por Cho et al. [Cho et al. 2014] objetivando capturar, adaptativamente, dependências em diferentes escalas de tempo. Similar à unidade da rede LSTM, a arquitetura GRU possui unidades de *gating*, que modulam o fluxo de informação dentro da unidade, porém, sem reter células de memória segregadas [Chung et al. 2014].

Seguindo a mesma metodologia adotada nos demais modelos, foi empregada uma camada com 416 neurônios na rede GRU no primeiro experimento, uma camada com 320 neurônios no segundo, e uma camada com 512 no terceiro. De forma similar tal como nos modelos prévios, adotou-se uma camada densa de saída na rede GRU, com um neurônio.

2.2.4. Busca de Hiperparâmetros e Métricas de Avaliação

Conforme descrito na seção anterior, foram aplicadas técnicas de ajuste de hiperparâmetros a fim de otimizar o desempenho de cada modelo de rede neural, buscando assim o melhor *fine-tuning* para cada implementação. De fato, a busca de hiperparâmetros desempenha um papel crucial no treinamento de modelos de aprendizado profundo, permitindo, desta forma, maximizar a acurácia dos modelos.

Para essa tarefa, foi empregada a biblioteca *Keras Tuner*, que oferece uma variedade de métodos de busca, sendo a otimização bayesiana uma das abordagens mais eficazes da literatura [Snoek et al. 2012]. Em termos mais técnicos, foram definidas 1.000 épocas, em adição à interrupção do processo de busca caso a função de *loss* do modelo não demonstrasse melhorias após 5 épocas. Os dados foram extraídos do conjunto de treinamento, sendo 20% desses reservados para validação e o restante para o treinamento. Vale ressaltar que o conjunto de teste foi mantido separado da busca e do treinamento, sendo usado apenas para realizar as previsões após o treinamento dos modelos.

A otimização bayesiana é uma técnica que busca maximizar ou minimizar uma função objetivo, neste caso a função $L(\theta|D)$ (função de perda *loss*) do modelo, através do uso do Teorema de Bayes. Essa abordagem se baseia na construção de um modelo probabilístico da função objetivo e na utilização de métodos estatísticos para determinar as próximas configurações de hiperparâmetros a serem avaliadas [Snoek et al. 2015].

Para a otimização bayesiana, considera-se uma distribuição a priori $P(\theta)$ sobre os hiperparâmetros θ [Frazier 2018]. Após verificação dos resultados, atualiza-se os hiperparâmetros através do Teorema de Bayes, obtendo a distribuição posterior $P(\theta|D)$:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)}, \quad \text{onde:}$$

- $P(\theta|D)$ é a distribuição posterior dos hiperparâmetros após observar os dados D .
- $P(D|\theta)$ é a verossimilhança dos dados D dado os hiperparâmetros θ .
- $P(\theta)$ é a distribuição priori dos hiperparâmetros.
- $P(D)$ é a probabilidade dos dados, atuando como uma constante de normalização.

Por fim, visando conduzir a análise quantitativa dos resultados gerados pelos modelos pós-treinados, foram computados o *Mean Absolute Percentage Error* (MAPE) e o *Root Mean Square Error* (RMSE) [Amaral et al. 2021a], a partir das formulas abaixo:

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \times 100, \text{ e } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}, \quad (1)$$

onde Y_i e \hat{Y}_i denota cada um dos n valores de referência e preditos.

3. Resultados e Discussão

Esta seção apresenta os resultados obtidos a partir da aplicação das redes pós-treinadas, bem como o design experimental dos testes, incluindo questões relativas às etapas de treinamento de cada teste.

3.1. Design dos Experimentos e Treinamento

Objetivando analisar o desempenho preditivo das redes, foram conduzidos três experimentos. No primeiro, foram utilizadas instâncias de 03/12/2022 à 31/01/2023, sendo 85% delas destinadas à construção do conjunto de treinamento, e 15% para o de teste. No segundo caso, foram coletados dados entre 03/12/2022 à 15/01/2023, com 80% das amostras usadas para treinamento, e 20% para teste. Em cada um dos experimentos, o objetivo foi verificar a robustez dos modelos frente a um cenário de declínio da doença, e de forte ascensão e declínio (um pico), respectivamente. Já o experimento final corresponde à janela de 03/12/2022 à 31/05/2023, de forma que 85% dos dados compõem o conjunto de treinamento, e 15% o de teste.

Após a realização de uma extensa bateria de testes, foi adotado 100 épocas para o treinamento, ponderando-se erro e custo moderado. O treinamento foi monitorado a partir da técnica *EarlyStopping*, interrompendo-se o processo se o erro não melhorasse após 10 épocas consecutivas. Nas três redes e em todos os experimentos, foi empregado a função de ativação *tanh* para transformação de entradas e estados ocultos, e a função *sigmoid* para camadas de saída. O otimizador Adam foi utilizado com uma taxa de aprendizado inicial de 0.001. Além disso, foram realizados testes variando-se a janela deslizante de 1 a 10, e de 14 a 21 dias. As janelas de 1 e 4 dias foram as que apresentaram os melhores resultados, justificando a escolha de 1 dia para a janela deslizante neste trabalho.

3.2. Experimentos

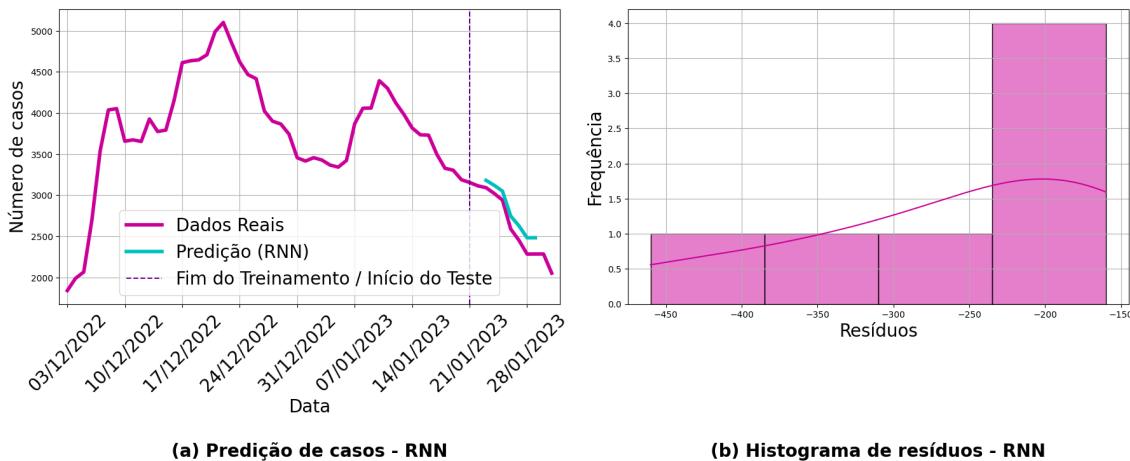
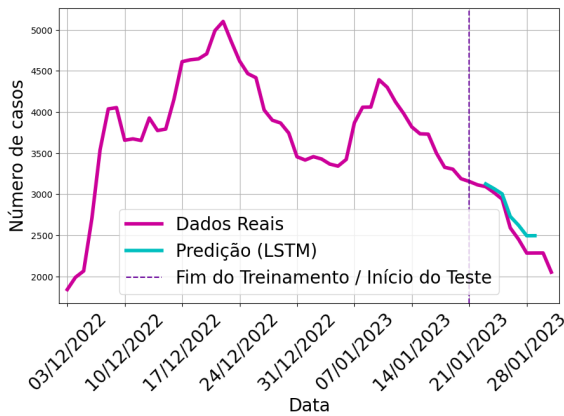
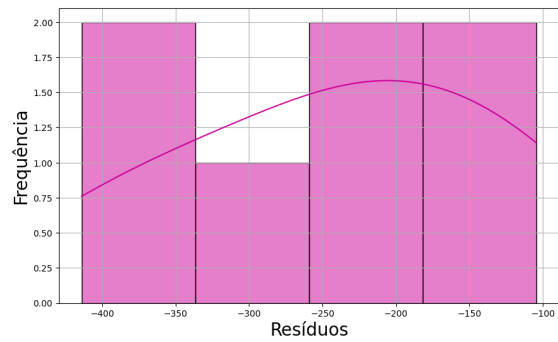


Figura 1. Predições para o modelo RNN e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 21/01/2023 à 31/01/2023.

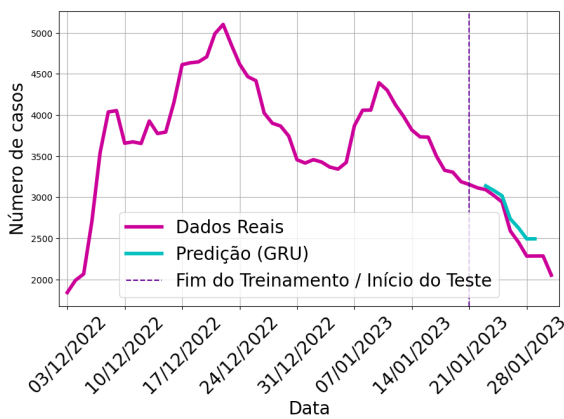


(a) Predição de casos - LSTM

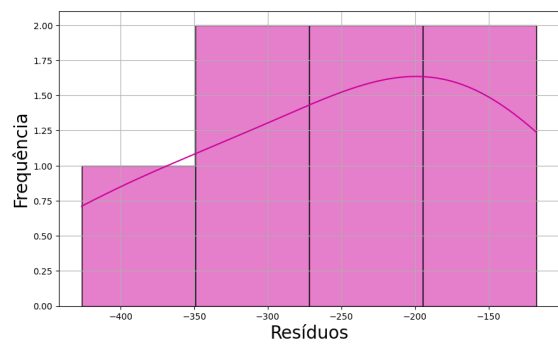


(b) Histograma de resíduos - LSTM

Figura 2. Predições para o modelo LSTM e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 21/01/2023 à 31/01/2023.



(a) Predição de casos - GRU



(b) Histograma de resíduos - GRU

Figura 3. Predições para o modelo GRU e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 21/01/2023 à 31/01/2023.

Nas Figuras 1(a), 2(a) e 3(a), são apresentados os gráficos das predições obtidas pelos modelos no primeiro experimento. Neste caso, os modelos foram submetidos a um período de queda nos casos diários de COVID-19. Conforme observado, os modelos capturaram com precisão tanto a tendência da curva como a magnitude em escala. Para uma análise mais aprofundada da qualidade dos treinamentos dos modelos, foram calculados os respectivos resíduos. Nas Figuras 1(b), 2(b) e 3(b), pode ser observado o histograma de resíduos para cada modelo. Apesar dos três modelos apresentarem predições acuradas, os histogramas que mais se assemelham a uma curva de distribuição normal são os dos modelos LSTM e GRU. Essa característica, de similaridade com a distribuição normal, é um indicativo que traduz uma maior acurácia conforme discutido em [Ma et al. 2015].

Para a realização do segundo experimento, foi selecionado um período cuja curva de casos apresentou um comportamento de pico, isto é, um aumento acentuado seguido de uma queda nos casos, com o objetivo de verificar se os modelos seriam capazes de prever mudança de tendência.

A partir das Figuras 4(a), 5(a) e 6(a), é possível verificar que os modelos foram capazes de capturar a assinatura da curva de casos de forma precisa, resultando em um desempenho bastante adequado para a tarefa de predição. Os histogramas de resíduos (Figuras 4(b), 5(b) e 6(b)) dos modelos apresentaram comportamentos similares.

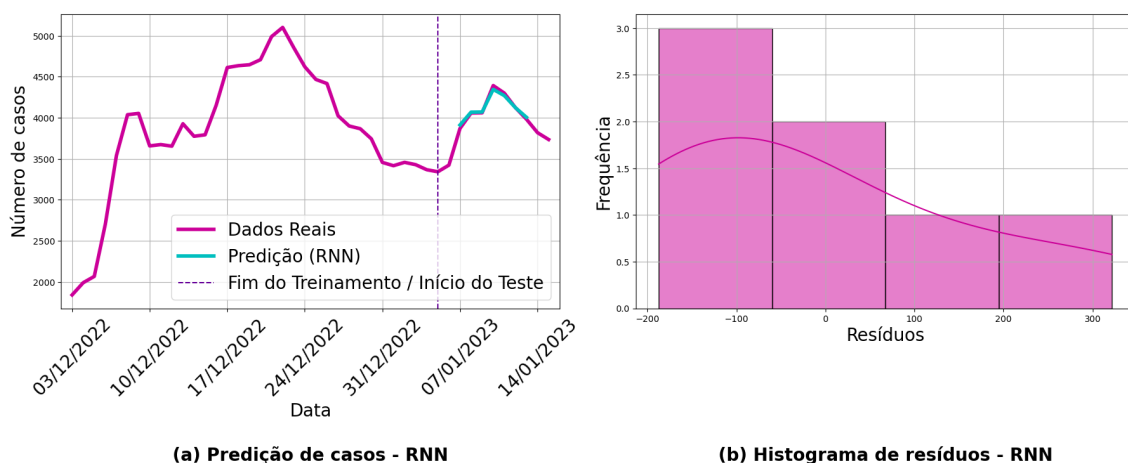


Figura 4. Predições para o modelo RNN e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 05/01/2023 à 15/01/2023.

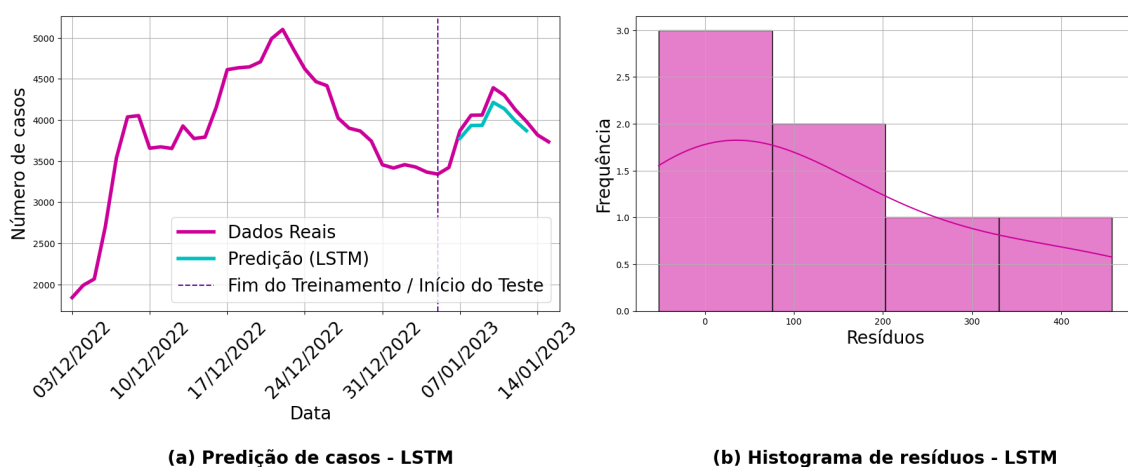


Figura 5. Predições para o modelo LSTM e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 05/01/2023 à 15/01/2023.

Finalmente, em nosso último experimento, selecionamos um período semelhante ao do experimento anterior, no qual a curva de casos apresentou um comportamento de pico, no entanto, ampliamos a janela de *forecasting* para 25 dias consecutivos, visando assim avaliar a capacidade dos modelos de prever sob um período de longo termo. Os resultados são apresentados na última coluna da Tabela 1.

Considerando as análises quantitativas das predições listadas na Tabela 1, que reúne os valores médios para o MAPE e RMSE, é possível verificar que os três modelos produziram medidas satisfatórias. Por exemplo, no experimento 1, os escores ficaram próximos de 10% para a métrica MAPE, enquanto que no experimento 2, a mesma métrica

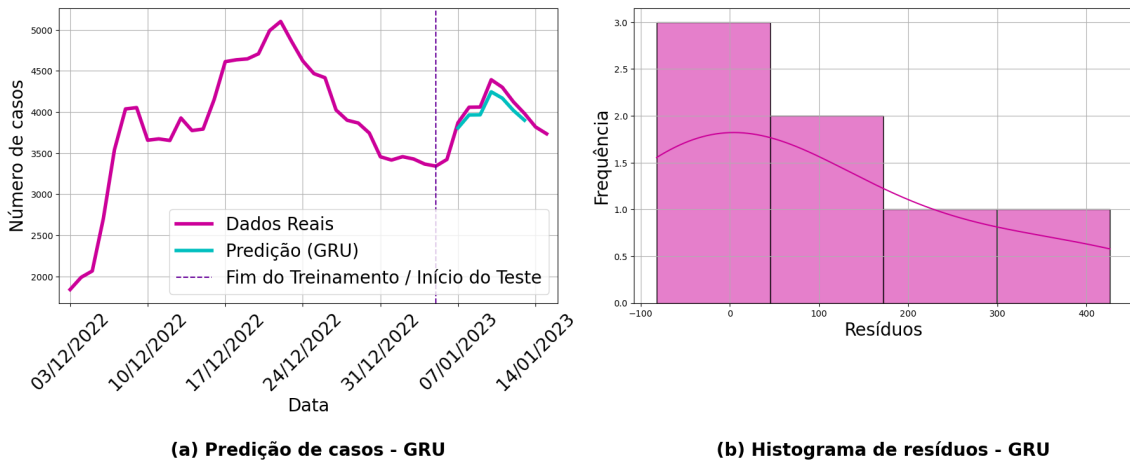


Figura 6. Predições para o modelo GRU e histograma de resíduos para o conjunto de teste, com casos de COVID-19 entre 05/01/2023 à 15/01/2023.

se manteve confinada na faixa de 3%. Já com relação ao experimento 3, a média do MAPE ficou ente 8% e 12%, o que evidencia a alta capacidade dos modelos em prever períodos inteiros de longo termo.

Tabela 1. MAPE e RMSE para as predições de casos da COVID-19.

Redes	Experimento 1		Experimento 2		Experimento 3	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
RNN	10.51%	280.99	3.46%	170.53	8.40%	153.42
GRU	10.00%	267.30	3.39%	195.61	10.18%	156.72
LSTM	9.76%	261.45	3.47%	211.99	12.29%	165.43

4. Considerações Finais e Trabalhos Futuros

Este trabalho apresentou uma avaliação de desempenho de três redes neurais recorrentes, a saber: *Recurrent Neural Network* (RNN), *Long Short-Term Memory* (LSTM), e *Gated Recurrent Unit* (GRU) para prever casos de COVID-19. Em contraste com outros estudos, que focam na análise da pandemia nos períodos mais severos da doença (2020 e 2021), nesta pesquisa procurou-se investigar o desempenho das referidas redes posto o contexto pandêmico atual, isto é, frente aos desafios capturar as tendências e escala das curvas a partir de dados com baixa recorrência de atualizações e ausência de ondas de grande magnitudes devido à vacinação massiva.

A partir da análise dos resultados, foi possível verificar que as três redes ajustadas apresentaram performance acurada para prever a dinâmica da doença, com validação tanto no cenário de “descida” da curva de casos como durante os recrudescimentos, tanto em períodos de 10 dias como de longo-termo – 25 dias. Considerando as métricas de avaliação, foi observado que no primeiro experimento, o MAPE se manteve próximo de 10%, enquanto que no experimento subsequente, tal métrica ficou abaixo de 3,5%, o que indica alta assertividade na predição. Já o experimento que considera 25 dias consecutivos de teste, este também apresentou escores MAPE entre 8% e 10%, o que novamente atesta a boa performance dos modelos frente ao desafio de prever a COVID-19.

Diante do exposto, este estudo procurou contribuir para o avanço de aplicações envolvendo redes neurais preditoras da COVID-19, sobretudo, no cenário atual da pandemia, no pós-vacinação. Considerando o potencial das redes neurais para predição e modelagem de séries temporais complexas, pretendemos utilizar os resultados deste estudo como base para investigações futuras com outras técnicas de aprendizado profundo. Neste caso, como trabalhos futuros, tem-se como vias a seguir: (a) utilizar dados de hospitalizações para o treinamento; (b) empregar dados de variantes da COVID-19 para melhor capturar dinâmicas de aumento; (c) adaptar outras técnicas de aprendizado profundo para prever os casos a partir dos dados de hospitalizações.

Agradecimentos

O presente trabalho foi financiado pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq – grants 164326/2020-0 e 316228/2021-4), a pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP – grant 2023/14427-8).

Referências

- Amaral, F., Casaca, W., Oishi, C. M., and Cuminato, J. A. (2021a). Simulating immunization campaigns and vaccine protection against COVID-19 pandemic in Brazil. *IEEE Access*, 9:126011–126022.
- Amaral, F., Casaca, W., Oishi, C. M., and Cuminato, J. A. (2021b). Towards providing effective data-driven responses to predict the covid-19 in São Paulo and Brazil. *Sensors*, 21(2):540.
- Capomaccio, S. (2024). Casos de covid-19 registram aumento e vacinação ainda é importante. *Jornal da USP*.
- Chandra, R., Jain, A., and Singh Chauhan, D. (2022). Deep learning via lstm models for covid-19 infection forecasting in india. *PloS one*, 17(1):e0262708.
- Cho, K., van Merriënboer, B., Bahdanau, D., and Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. In Wu, D., Carpuat, M., Carreras, X., and Vecchi, E. M., editors, *Proceedings of SSST-8, Eighth Workshop on Syntax, Semantics and Structure in Statistical Translation*, pages 103–111, Doha, Qatar. Association for Computational Linguistics.
- Chung, J., Gulcehre, C., Cho, K., and Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*.
- Davahli, M., Fiok, K., Karwowski, W., Aljuaid, A., and Taiar, R. (2021). Predicting the dynamics of the covid-19 in the united states using graph theory-based neural networks. *International Journal of Environmental Research and Public Health*, 18(7).
- Frazier, P. I. (2018). A tutorial on bayesian optimization. *arXiv preprint arXiv:1807.02811*.
- Gomez-Cravioto, D. A., Diaz-Ramos, R. E., Cantu-Ortiz, F. J., and Ceballos, H. G. (2021). Data analysis and forecasting of the covid-19 spread: A comparison of recurrent neural networks and time series models. *Cognitive Computation*, pages 1–12.

- Hawas, M. (2020). Generated time-series prediction data of covid-19's daily infections in brazil by using recurrent neural networks. *Data in Brief*, 32:106175.
- Hewamalage, H., Bergmeir, C., and Bandara, K. (2021). Recurrent neural networks for time series forecasting: Current status and future directions. *International Journal of Forecasting*, 37(1):388–427.
- Linge, S. and Langtangen, H. P. (2020). *Programming for Computations-Python: A Gentle Introduction to Numerical Simulations with Python 3.6*. Springer Nature.
- Ma, J., Zhao, H., and Schuck, P. (2015). A histogram approach to the quality of fit in sedimentation velocity analyses. *Analytical Biochemistry*, 483:1–3.
- Mendes, V. (2024). Por que covid-19 ainda mata tanta gente no brasil. BBC News Brasil.
- Microsoft (2023). Microsoft. <https://powerbi.microsoft.com/pt-br/what-is-power-bi/>. Acesso em 20 de julho de 2023.
- Ministério da Saúde (2024). Relatório da situação de saúde diante da pandemia de covid-19. Disponível em: https://bvsms.saude.gov.br/bvs/publicacoes/saude_brasil_2020-2021_analise_pandemia_covid-19.pdf. Acesso em: 14 mar. 2024.
- Oliveira, C. (2024). Controle sobre a covid-19 ainda é instável no país após quatro anos do 1º caso. Brasil de Fato.
- Roosa, K., Lee, Y., Luo, R., Kirpich, A., Rothenberg, R., Hyman, J. M., Yan, P., and Chowell, G. (2020). Short-term forecasts of the COVID-19 epidemic in Guangdong and Zhejiang, China: February 13–23, 2020. *Journal of Clinical Medicine*, 9:596.
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015). Convolutional lstm network: A machine learning approach for precipitation nowcasting. *Advances in neural information processing systems*, 28.
- Snoek, J., Larochelle, H., and Adams, R. P. (2012). Practical bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc.
- Snoek, J., Rippel, O., Swersky, K., Kiros, R., Satish, N., Sundaram, N., Patwary, M., Prabhat, M., and Adams, R. (2015). Scalable bayesian optimization using deep neural networks. In *International conference on machine learning*, pages 2171–2180. PMLR.
- Worldometer (2024). Covid-19 Coronavirus pandemic. Disponível em: <https://www.worldometers.info/coronavirus>. Acesso em: 14 mar. 2024.
- Worobey, M. (2021). Dissecting the early COVID-19 cases in Wuhan. *Science*, 374(6572):1202–1204.
- Zeroual, A., Harrou, F., Dairi, A., and Sun, Y. (2020). Deep learning methods for forecasting covid-19 time-series data: A comparative study. *Chaos, Solitons & Fractals*, 140:110121.