

Um Modelo Profundo Crescente para Previsão de Óbitos por COVID-19 no Brasil

¹Ricardo de A. Araújo, ¹Eliseu C. de Brito, ²Ya-Sin B. Mghazli, ³Sergio Soares e ²José M. de Seixas

¹Laboratório de Inteligência Computacional do Araripe, Instituto Federal do Sertão Pernambucano, Ouricuri, PE, Brasil.

²Laboratório de Processamento de Sinais, COPPE/POLI, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.

³Centro de Informática, Universidade Federal de Pernambuco, Recife, PE, Brasil.

ricardo.araujo@ifsertao-pe.edu.br, eliseu.cordeiro@aluno.ifsertao-pe.edu.br, yasin.barcelos@lps.ufrj.br

scbs@cin.ufpe.br, seixas@lps.ufrj.br

Resumo—A doença do coronavírus 2019 (COVID-19) foi declarada pela Organização Mundial da Saúde (OMS) como uma pandemia sem precedentes, tendo sobrecarregado os sistemas de saúde em todo o mundo devido à alta demanda de internações em unidades de terapia intensiva. Nesse cenário, entender a dinâmica da pandemia é essencial para orientar a estratégia para lidar simultaneamente com o aumento da demanda por apoio à saúde e os efeitos diretos e colaterais na economia. Neste trabalho é apresentado um estudo empírico sobre a série temporal de óbitos por COVID-19 no Brasil e, baseado neste estudo, propomos um modelo de dilatação-erosão profunda, projetado por um processo de aprendizado baseado em gradiente descendente, capaz de prever esse tipo particular de série temporal. Os resultados obtidos mostram que o modelo proposto supera modelos clássicos e recentes de aprendizado de máquina apresentados na literatura para estimar a dinâmica da pandemia de COVID-19 no Brasil.

Index Terms—Rede Neural Profunda, Operador de Dilatação-Erosão, Aprendizagem baseada em Gradiente Descendente, Previsão de Séries Temporais, COVID-19.

I. INTRODUÇÃO

Atualmente, o controle de doenças virais ainda é considerado um grande desafio para a saúde pública mundial [1]. Nesse sentido, os coronavírus pertencem à classe dos vírus de ácido ribonucléico envelopado que têm sido responsáveis por várias doenças no sistema respiratório humano [2].

Devido à sua alta diversidade genética, prevalência e recombinação de genes, um novo coronavírus, denominado *severe acute respiratory syndrome coronavirus 2* (SARS-Cov-2) [3], surgiu em Wuhan - Hubei (China), em meados de 2019, originado a partir de morcegos [4]. O SARS-Cov-2 é o agente etiológico da doença do coronavírus 2019 (COVID-19) [5], onde na maioria dos casos confirmados são observados sintomas leves [6].

No entanto, baseado em evidências relacionadas a uma possível disseminação acelerada, o COVID-19 foi reconhecido como uma pandemia pela Organização Mundial de Saúde [7] em março de 2020. A pandemia se espalhou pelo mundo, desencadeando uma crise catastrófica sem precedentes [8], que se tornou uma ameaça real à saúde e à economia global [9].

Tentando evitar o colapso dos sistemas de saúde, chefes de Estado empregaram medidas de quarentena, fechamento de escolas, distanciamento social, uso de máscaras e restrições de viagens para retardar a disseminação acelerada do COVID-19 [10], uma vez que não havia terapia medicamentosa eficaz e vacinas licenciadas para prevenir a COVID-19 [11]. No entanto, essas medidas tiveram impacto em diversas atividades econômicas, causando efeitos negativos em todo o mundo [8].

Além disso, o surgimento de variantes do SARS-Cov-2 tem sido associado a complicações moderadas em torno de 15-20% dos pacientes e complicações graves em torno de 5-12% dos pacientes, havendo uma sobrecarga além do esperado nos sistemas de saúde globais [12], mesmo com a rápida implantação de vacinas com autorização de uso emergencial [13].

Nesse contexto, surgiram preocupações sobre a eficácia e os efeitos colaterais das vacinas ao longo do processo de vacinação em todo o mundo [14]. Muitas evidências comprovaram sua eficácia contra complicações moderadas e graves da COVID-19 [15].

No entanto, existem algumas questões em aberto sobre a diminuição da imunidade e proteção reduzida das vacinas contra variantes do SARS-Cov-2 [15]–[17], como a variante Alpha (com pico por volta de março-abril de 2021), a variante Delta (com pico por volta de agosto de 2021), a variante Omicron (com pico por volta do início de 2022) e, mais recentemente, as subvariantes BQ.1 e BQ.1.1 Omicron (no segundo semestre de 2022, principalmente nos Estados Unidos e no Brasil).

Nesse sentido, vários esforços contra os efeitos diretos e colaterais do SARS-Cov-2 (como o fator de proteção e a caracterização da duração da imunidade) continuam mesmo depois do início da pandemia [18], [19]. Nesse cenário, prever a dinâmica de uma pandemia é essencial para subsidiar ações estratégicas para serviços de saúde e economia [9], [20], [21].

Vários modelos epidemiológicos podem ser encontrados na literatura tentando prever a pandemia da COVID-19 [22]–[28].

No entanto, Roda *et al.* [29] e Benvenuto *et al.* [30] argumentaram que esses modelos não obtiveram sucesso para prever a pandemia da COVID-19 devido a necessidade de intervenção humana (especialistas no problema) para a definição dos parâmetros dos modelos de forma a garantir um desempenho minimamente aceitável. Tentando superar essa desvantagem, o modelo *autoregressive integrated moving average* (ARIMA) [31], [32] foi investigado para prever séries temporais da COVID-19. No entanto, como um modelo linear típico, tem uma desvantagem intrínseca para estimar o fenômeno gerador das séries temporais da COVID-19, uma vez que estas possuem algum tipo de não-linearidade em seu fenômeno gerador [33].

Por esse motivo, modelos de *machine learning* (ML) [18] têm sido considerados alternativas promissoras para construir mapeamentos capazes de estimar o futuro da pandemia da COVID-19 [34]. Nesse contexto, podemos encontrar trabalhos utilizando *multilayer perceptron* (MLP) [35], *radial basis function neural network* (RBF) [36], *nonlinear autoregressive neural network with exogenous inputs* (NARX) [31], [36], *support vector regressor* (SVR) [35], [37], *long short-term memory neural network* (LSTM) [31], [32], [37], *convolutional neural network* (CNN) [37], entre outros.

Dessa forma, tentando responder a algumas questões sobre a dinâmica dessa pandemia, é apresentado um estudo sobre o fenômeno gerador da série temporal de óbitos por COVID-19 no Brasil. Com base neste estudo, propomos um modelo de dilatação-erosão profunda projetado por um processo de aprendizado baseado em gradiente descendente para prever este tipo particular de série temporal. O *mean squared error* (MSE) é usado para avaliar o desempenho da previsão, que é validado estatisticamente usando os testes de Friedman e Tukey. Os resultados alcançados com o modelo proposto mostram um desempenho superior quando comparado aos modelos clássicos e recentes da literatura para estimar a pandemia da COVID-19.

Este trabalho está organizado da seguinte forma. Na Seção II, é apresentado um estudo empírico da série temporal de óbitos por COVID-19 no Brasil. Em seguida, na Seção III, é descrito o modelo proposto. Posteriormente, na Seção IV, é delineada a metodologia de simulação, bem como são descritos os resultados experimentais. Ao final, na Seção V, são apresentadas algumas conclusões úteis e são apontadas algumas direções futuras para este trabalho.

II. ANÁLISE DA SÉRIE TEMPORAL

Uma série temporal relacionada a casos de óbito por COVID-19 no Brasil (BR) é investigada neste trabalho, que teve foco nessa série porque o Brasil teve surtos significativos de COVID-19, com um alto número de casos e óbitos relatados, o que pode fornecer várias informações sobre a dinâmica

da pandemia em regiões que enfrentaram grandes desafios. Além disso, a disponibilidade de seus dados permitiram a captura de uma gama mais ampla de cenários, possibilitando uma compreensão mais abrangente dos fatores que influenciam a disseminação e o impacto da COVID-19. A série temporal é amostrada em frequência diária (no período de 01/04/2020 a 31/07/2022), conforme ilustrado na Figura 1.

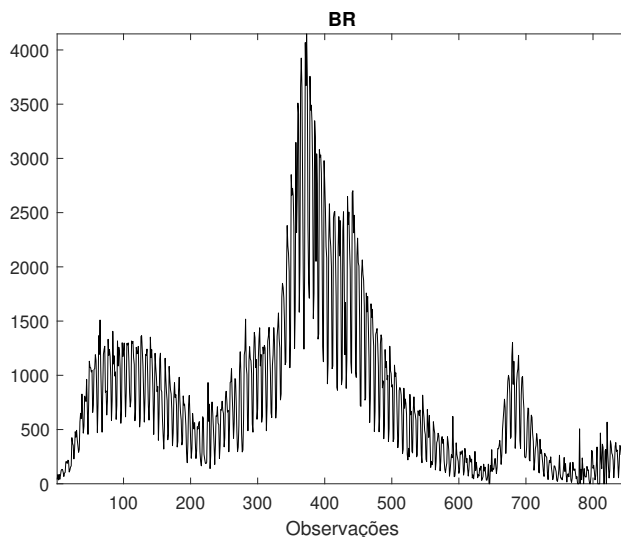


Figura 1. Gráfico da série temporal.

Em seguida, é analisada a função de autocorrelação (ACF) [38] e a função de autocorrelação parcial (PACF) [38], representadas nas Figuras 2 e 3, respectivamente. Observe a presença de um decaimento lento com ciclos regulares nas curvas da ACF. As curvas da PACF são caracterizadas por uma alta correlação em defasagens de baixa ordem, que diminui com o aumento na ordem dos retardos. Portanto, a análise da ACF e PACF sugere a presença de dependência linear e não-linear na série temporal investigada.

No entanto, tal análise não permite uma avaliação correta da dependência não-linear. Dessa forma, é analisada a informação mútua média (MMI) [39], representada na Figura 4, onde também é possível observar um decaimento lento com ciclos regulares, confirmando a presença de uma relação não-linear.

Em seguida, é investigado o parâmetro de Hurst (HP) [40], representado na Figura 5, para avaliar a natureza da relação não-linear encontrada na análise da MMI. Note que o HP está próximo de 0 para os retardos temporais considerados, tendo um comportamento anti-persistente, ou seja, um processo auto-semelhante com dependência não-linear de longo prazo. Portanto, é possível concluir que existe algum tipo de relação entre previsibilidade e a escolha dos retardos temporais.

De acordo com a análise previamente apresentada, parece razoável assumir que o fenômeno gerador da série temporal investigada pode ser modelado por uma combinação de uma

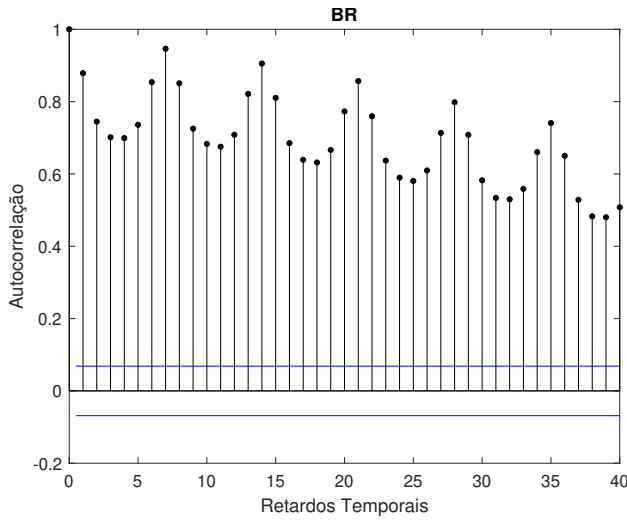


Figura 2. Função de autocorrelação da série temporal.

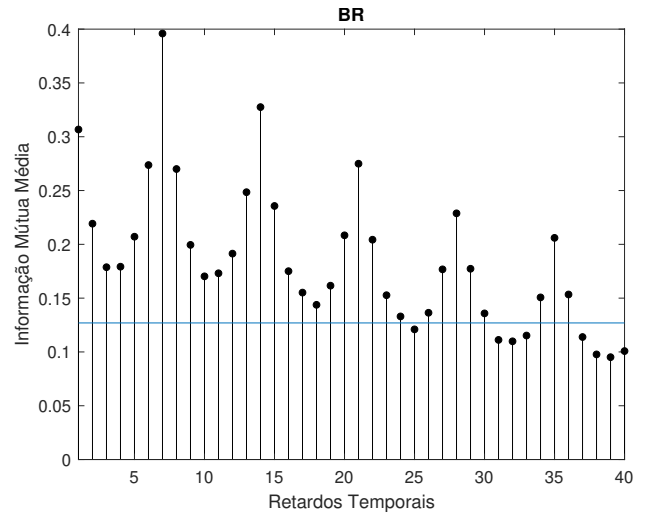


Figura 4. Informação mútua média da série temporal.

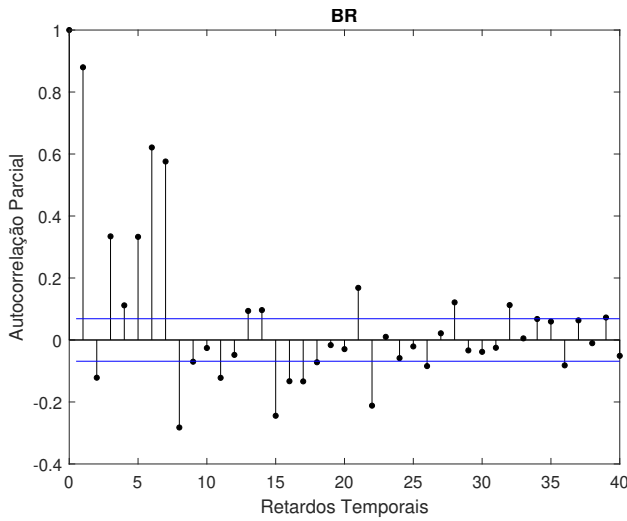


Figura 3. Função de autocorrelação parcial da série temporal.

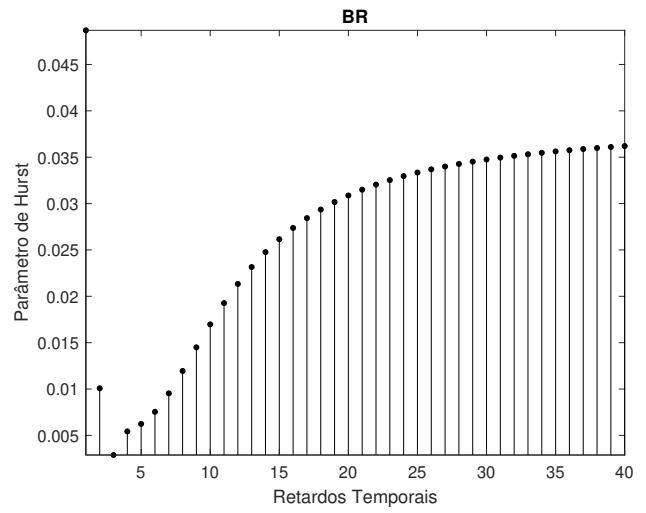


Figura 5. Parâmetro de Hurst da série temporal.

componente linear e uma componente não-linear. Neste contexto, seja $\Psi : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função crescente, na qual n é a dimensionalidade dos retardos temporais, e a seguinte relação, para todo $l, s \in \{n + 1, \dots, N\}$:

$$(x_{l-n}, \dots, x_{l-1}) \leq (x_{s-n}, \dots, x_{s-1}) \Rightarrow x_l \leq x_s. \quad (1)$$

É possível verificar que a série temporal analisada satisfaz a Equação 1 para os retardos temporais considerados nos experimentos realizados na Seção IV. Portanto, assumimos implicitamente que a componente não-linear pode ser estimada em termos de funções crescentes.

III. O MODELO PROPOSTO

A motivação deste trabalho está relacionada ao fato dos modelos epidemiológicos não terem sido bem sucedidos na previsão da pandemia da COVID-19 [29]. Nesse sentido,

a dependência humana para definir esses parâmetros epidemiológicos compromete o projeto automatizado de um modelo com desempenho minimamente aceitável [30]. Ao contrário dos modelos epidemiológicos, este trabalho emprega uma abordagem baseada em séries temporais para estimar cenários futuros da pandemia de COVID-19. Além disso, considerando que Araújo *et al.* [41] demonstrou que uma função crescente pode ser aproximada usando operadores morfológicos de dilatação e erosão, propomos um modelo profundo, chamado de *deep dilation-erosion neural network* (DeepDE). A Figura 6 apresenta a arquitetura do modelo proposto.

De acordo com a Figura 6, a i -ésima saída de uma unidade de processamento (DE) na j -ésima camada de dilatação-erosão da rede neural DeepDE é dada por :

$$y_i^{(j)} = f(u_i^{(j)}), \quad (2)$$

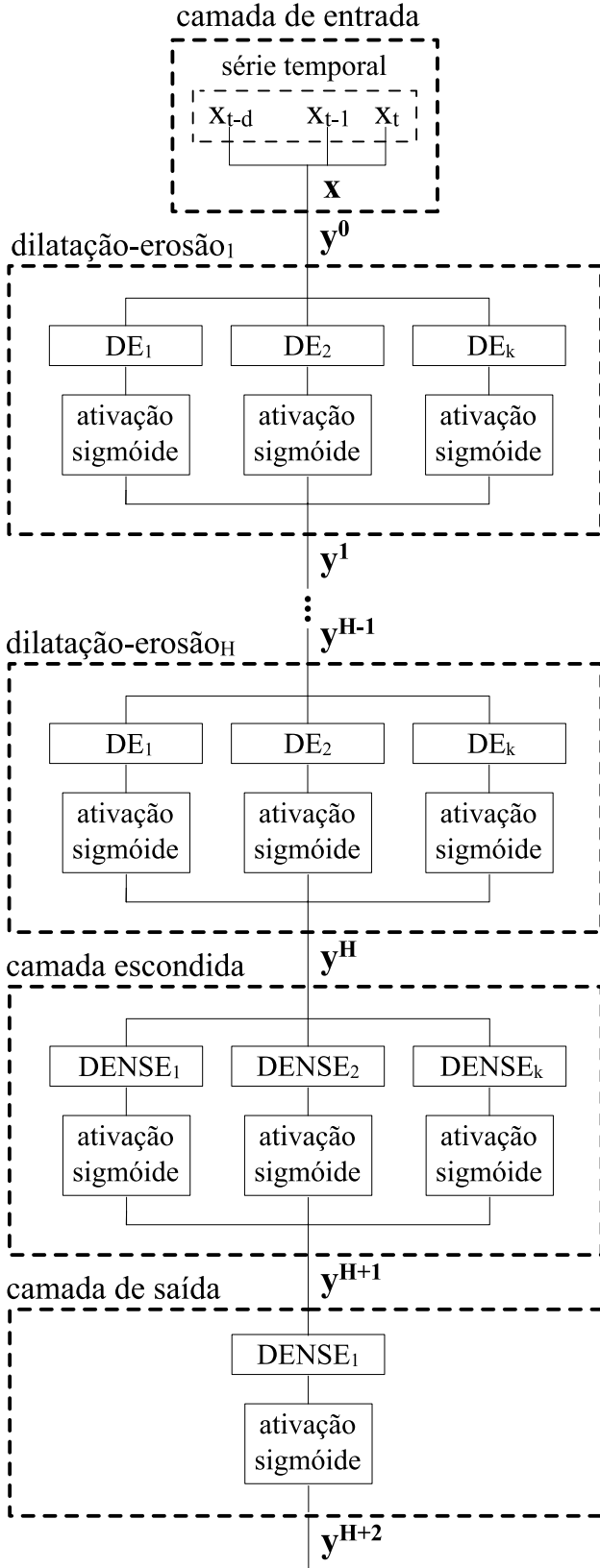


Figura 6. Arquitetura do modelo proposto.

onde

$$u_i^{(j)} = \lambda_i^{(j)} \delta_i^{(j)} + (1 - \lambda_i^{(j)}) \varepsilon_i^{(j)}, \quad (3)$$

com

$$\delta_i^{(j)} = \delta_{\mathbf{a}_i^{(j)}}(\mathbf{y}^{(j-1)}) = \bigvee_{k=1}^{I_{j-1}} (y_k^{(j-1)} + a_{i,k}^{(j)}), \quad (4)$$

e

$$\varepsilon_i^{(j)} = \varepsilon_{\mathbf{b}_i^{(j)}}(\mathbf{y}^{(j-1)}) = \bigwedge_{k=1}^{I_{j-1}} (y_k^{(j-1)} + b_{i,k}^{(j)}), \quad (5)$$

em que I_j representa a quantidade de unidades de processamento na j -ésima camada, $\lambda_i^{(j)} \in \mathbb{R}$, e $\mathbf{a}_i^{(j)}, \mathbf{b}_i^{(j)} \in \mathbb{R}^{I_{j-1}}$. A ativação interna ($u_i^{(j)}$) é dada por uma combinação linear (definida por $\lambda_i^{(j)}$) entre um operador de dilatação ($\delta_i^{(j)}$) e um operador de erosão ($\varepsilon_i^{(j)}$). Os termos \mathbf{a} e \mathbf{b} são os elementos estruturantes da dilatação ($\delta_{\mathbf{a}_i^{(j)}}(\mathbf{y}^{(j-1)})$) e da erosão ($\varepsilon_{\mathbf{b}_i^{(j)}}(\mathbf{y}^{(j-1)})$), respectivamente.

Também, de acordo com a Figura 6, a i -ésima saída da unidade de processamento (DENSE) na j -ésima camada da rede neural DeepDE é dada por:

$$y_i^{(j)} = f(u_i^{(j)}), \quad (6)$$

onde

$$u_i^{(j)} = \sum_{k=1}^{I_{j-1}} y_k^{(j-1)} c_{i,k}^{(j)} + \rho_i^{(j)}, \quad (7)$$

em que $\rho_i^{(j)} \in \mathbb{R}$, e $\mathbf{c}_i^{(j)} \in \mathbb{R}^{I_{j-1}}$.

Note que qualquer função de ativação pode ser utilizada nessas unidades de processamento. Optou-se pela função de ativação sigmóide (uma vez que os operadores morfológicos de dilatação e erosão utilizados no modelo proposto devem ser empregados entre reticulados completos, que é o caso do intervalo unitário $[0, 1]$), que é dada por:

$$f(u_i^{(j)}) = \frac{1}{1 + \exp(-u_i^{(j)})}. \quad (8)$$

Uma propriedade importante de uma rede neural é sua capacidade de aprender usando um processo iterativo para ajustar seus pesos sinápticos [42]. Tal processo de aprendizagem é caracterizado pela presença de um agente externo que induz a rede neural a uma resposta desejada, a fim de construir um mapeamento entre entradas e saídas por meio da minimização de uma função de custo, de forma que a resposta observada se aproxime da resposta desejada a cada iteração, definida como época, no processo de aprendizagem [42].

A função de custo define uma superfície de erro sobre o espaço de pesos. O método de otimização usado para minimizar a função de custo usa informações do gradiente descendente para ajustar os parâmetros da rede neural. Esses métodos encontram pontos de mínimo (locais ou globais) na superfície de erro a partir de uma condição inicial arbitrária. O

método clássico utilizado no processo de aprendizado de redes neurais, que utiliza informações do gradiente descendente, é o algoritmo *backpropagation* (BP) [42], e, por esse motivo, foi escolhido para treinar o modelo proposto. Este requer o ajuste do vetor pesos (\mathbf{w}), definido por:

$$\mathbf{w}_i^{(j)} = (\lambda_i^{(j)}, \mathbf{a}_i^{(j)}, \mathbf{b}_i^{(j)}) \text{ para unidades DE,} \quad (9)$$

ou

$$\mathbf{w}_i^{(j)} = (\mathbf{c}_i^{(j)}, \rho_i^{(j)}), \text{ para unidades DENSE,} \quad (10)$$

em que $i = 1, \dots, I_j$ com I_j representa a quantidade de unidades de processamento na j -ésima camada, e $j = 1, 2$.

O ajuste do vetor de pesos utiliza um critério de erro em termos da função de custo C , dada por:

$$C(i) = \frac{1}{M} \sum_{m=1}^M \xi(m), \quad (11)$$

em que

$$\xi(m) = \|\mathbf{e}(m)\|^2 \quad (12)$$

com

$$\|\mathbf{e}(m)\|^2 = \sum_{i=1}^{I_j} [e_i(m)]^2, \quad (13)$$

onde M é o número de amostras de treinamento, e $e_i(m)$ é o i -ésimo erro instantâneo para a m -ésima amostra de treinamento, dado por:

$$e_i(m) = t_i(m) - y_i^{(j)}(m), \quad (14)$$

em que $t_i(m)$ e $y_i^{(j)}(m)$ são as i -ésimas saídas desejadas e previstas, respectivamente.

O problema no ajuste do vetor de pesos é encontrar um ponto de mínimo na superfície da função de custo. Logo, o vetor de pesos pode ser estimado por:

$$\mathbf{w}_i^{(j)}(m+1) = \mathbf{w}_i^{(j)}(m) + \mu \nabla C_i^{(j)}(m), \quad (15)$$

onde μ é a taxa de aprendizagem e $\nabla C_i^{(j)}$ é o gradiente da função de custo, dado por:

$$\nabla C_i^{(j)}(m) = \frac{\partial \xi(m)}{\partial \mathbf{w}_i^{(j)}}. \quad (16)$$

Dessa forma, $\frac{\partial \xi(i)}{\partial \mathbf{w}_i^{(j)}}$ pode ser avaliado utilizando a regra da cadeia:

$$\frac{\partial \xi}{\partial \mathbf{w}_i^{(j)}} = \frac{\partial \xi}{\partial y_i^{(j)}} \frac{\partial y_i^{(j)}}{\partial u_i^{(j)}} \frac{\partial u_i^{(j)}}{\partial \mathbf{w}_i^{(j)}}. \quad (17)$$

A derivada parcial $\frac{\partial \xi}{\partial y_i^{(j)}}$ é dada por:

$$\frac{\partial \xi}{\partial y_i^{(j)}} = e_i, \quad \text{quando } j = J, \quad (18)$$

e

$$\frac{\partial \xi}{\partial y_i^{(j)}} = \sum_{i=1}^{I_{j+1}} \frac{\partial \xi}{\partial y_i^{(j+1)}} \frac{\partial y_i^{(j+1)}}{\partial y_i^{(j)}}, \quad \text{quando } j < J, \quad (19)$$

em que

$$\frac{\partial \xi}{\partial y_i^{(j+1)}} = e_i^{(j+1)}, \quad (20)$$

e

$$\frac{\partial y_i^{(j+1)}}{\partial y_i^{(j)}} = f(u_i^{(j+1)}) \frac{\partial u_i^{(j)}}{\partial y_i^{(j-1)}}. \quad (21)$$

A derivada parcial $\frac{\partial y_i^{(j)}}{\partial u_i^{(j)}}$ é dada por:

$$\frac{\partial y_i^{(j)}}{\partial u_i^{(j)}} = f'(u_i^{(j)}) \quad (22)$$

A derivada parcial $\frac{\partial u_i^{(j)}}{\partial \mathbf{w}_i^{(j)}}$ é dada por:

$$\frac{\partial u_i^{(j)}}{\partial \mathbf{w}_i^{(j)}} = \left(\frac{\partial u_i^{(j)}}{\partial \lambda_i^{(j)}}, \frac{\partial u_i^{(j)}}{\partial \mathbf{a}_i^{(j)}}, \frac{\partial u_i^{(j)}}{\partial \mathbf{b}_i^{(j)}} \right) \text{ para unidades DE,} \quad (23)$$

ou

$$\frac{\partial u_i^{(j)}}{\partial \mathbf{w}_i^{(j)}} = \left(\frac{\partial u_i^{(j)}}{\partial \mathbf{c}_i^{(j)}}, \frac{\partial u_i^{(j)}}{\partial \rho_i^{(j)}} \right) \text{ para unidade DENSE.} \quad (24)$$

A avaliação das derivadas $\frac{\partial u_i^{(j)}}{\partial \mathbf{y}^{(j-1)}}$, f' , $\frac{\partial u_i^{(j)}}{\partial \mathbf{c}_i^{(j)}}$, $\frac{\partial u_i^{(j)}}{\partial \rho_i^{(j)}}$ e $\frac{\partial u_i^{(j)}}{\partial \lambda_i^{(j)}}$ é feita utilizando a metodologia apresentada em [41].

IV. SIMULAÇÕES E RESULTADOS EXPERIMENTAIS

Em todas as simulações realizadas, a série temporal foi normalizada para o intervalo $[0, 2, 0, 8]$ e dividimos os dados em três conjuntos. Para o conjunto de teste, foram utilizadas as últimas 31 observações (de 01/07/2022 a 31/07/2022). Nas observações restantes foi utilizado 90% dos dados para o conjunto de treinamento e 10% dos dados para o conjunto de validação.

Os resultados obtidos com o modelo proposto são comparados com os obtidos por modelos clássicos e relevantes apresentados na literatura: i) modelos estatísticos (ARIMA [30], LINREG [30], REGTREE [43] e SVR [44]), ii) modelos de rede neural (MLP [35], RBF [36], CNN [37]), e iii) modelos dinâmicos (NARX [31] e LSTM [32]).

Para os experimentos com os modelos ARIMA e REGTREE, foi empregada a metodologia sugerida em [30] e [43], respectivamente, utilizando a *econometrics toolbox* do *MATLAB*. Para os experimentos com o modelo SVR, foi empregada a metodologia sugerida por [44], utilizando a biblioteca *scikit-learn* do *Python*. Para os experimentos com modelos MLP, RBF e NARX, foram empregados os procedimentos apresentados em [35], [36] e [31], respectivamente, também usando a *neural network toolbox* do *MATLAB*. Para os experimentos com modelos LSTM e CNN, foi empregado o procedimento descrito em [32] e [37], respectivamente, usando a biblioteca *keras* do *Python*.

Definimos a arquitetura DeepDE ($I; \mathbf{H} = h_1, \dots, h_k; D_k, O; \mu; \sigma$) para os experimentos com o modelo

proposto. O termo I é a dimensionalidade de entrada, h_k é o número de unidades de processamento na k -ésima camada de dilatação-erosão, D_k é o número de unidades de processamento da camada oculta DENSE, O é a dimensionalidade de saída ($O=1$ devido ao foco deste trabalho na previsão de um-passo-à-frente), μ é a taxa de aprendizagem e σ é o fator de suavização. Para o número de camadas de dilatação-erosão (k), utilizamos a validação cruzada, onde são investigados os valores 1, 3 e 5. Fixamos o número de unidades de processamento ($h_k=10$) unidades para cada camada de dilatação-erosão do modelo. Também usamos validação cruzada para determinar o número de unidades de processamento DENSE (D_k), onde os valores 10, 50 e 100 são investigados.

Para a taxa de aprendizado (μ), utilizamos também a validação cruzada, onde são investigados os valores 0.001, 0.01 e 0.1. Para o fator de suavização (σ), empregamos a validação cruzada, onde são investigados os valores 0.005, 0.05 e 0.5. Definimos os valores iniciais $\mathbf{a}_i^{(j)}$, $\mathbf{b}_i^{(j)}$, $\mathbf{c}_i^{(j)}$, $\rho_i^{(j)} \in [0, 1]$ e $\lambda_i^{(j)} \in [0, 1]$. Três condições de parada são usadas para o processo de aprendizado: i) épocas de treinamento ($epoch = 10000$), ii) *process training* ($Pt \leq 10^{-6}$) e iii) *generalization loss* ($Gl > 5\%$). Além disso, para todos os modelos, a escolha da dimensionalidade de entrada é baseada na análise da série temporal apresentada na Seção II e dada por 1-9, 11, 13-15. Vale ressaltar que o modelo proposto foi desenvolvido utilizando o *MATLAB*.

Para avaliar o desempenho da previsão, usamos o *mean squared error* (MSE), dado por:

$$MSE = \frac{1}{N} \sum_{i=1}^N (z_t - \hat{z}_t)^2, \quad (25)$$

em que N é o número de amostras de série temporal e z_t e \hat{z}_t são os valores reais e previstos, respectivamente, para a t -ésima amostra da série temporal.

Para cada configuração do modelo, foram realizados trinta experimentos, onde são calculados a média (MEAN) e o desvio padrão (STD). Aplicamos também o teste de Friedman [45] com nível de significância $\alpha = 0.05$, pois estabelece um ranqueamento para todos os modelos. Além disso, usamos um teste *post hoc*, chamado teste de Tukey [46] com $\alpha = 0.05$, para avaliar o desempenho par a par entre os modelos.

Um resumo dos resultados para a série temporal investigada é apresentado na Tabela I, de acordo com as estatísticas MEAN e STD para a medida MSE. Vale ressaltar que o melhor desempenho é alcançado pelo modelo proposto (DeepDE), onde os valores obtidos da medida MSE indicam que a previsão está bastante próxima aos valores reais da série temporal.

O resultado do teste de Friedman para a medida MSE é apresentado na Tabela II, onde é possível confirmar estatisticamente os resultados apresentados na Tabela I. O modelo

Tabela I
DESEMPENHO NO CONJUNTO DE TESTE PARA A MEDIDA MSE.

Modelo	MSE
ARIMA	2.7306e-04 (0.0000e+00)
CNN	2.5933e-04 (7.3026e-05)
DeepDE	2.0363e-04 (3.1521e-05)
LINREG	3.4159e-04 (0.0000e+00)
LSTM	2.7827e-04 (8.7904e-05)
MLP	3.1371e-04 (2.6481e-08)
NARX	2.4495e-04 (1.0260e-04)
RBF	3.2225e-04 (0.0000e+00)
REGTREE	4.1009e-04 (0.0000e+00)
SVR	2.3797e-04 (0.0000e+00)

proposto obteve o menor valor de *rank*, indicando que é o modelo mais preciso para a série temporal investigada, considerando a medida MSE.

Tabela II
RESULTADO DO TESTE DE FRIEDMAN PARA A MEDIDA MSE.

Posição	$\chi^2=33.49$ and $p\text{-value}=5.01e-05$	
	Modelo	Rank
1	DeepDE	2.20
2	SVR	3.40
3	NARX	3.80
4	CNN	4.40
5	ARIMA	4.80
6	LSTM	5.20
7	MLP	6.00
8	RBF	7.00
9	LINREG	8.20
10	REGTREE	10.00

O resultado do teste de Tukey para a medida MSE é apresentado na Tabela III, sugerindo que o modelo proposto apresenta melhor desempenho em relação a todos os pares. O menor valor para a estatística do teste de Tukey é 1.20 (em relação ao par DeepDE - SVR), significando que o modelo proposto tem, estatisticamente, melhor desempenho considerando melhor dentre os investigados.

Tabela III
RESULTADO DO TESTE DE TUKEY PARA A MEDIDA MSE.

Pair	Statistic	p-value
DeepDE - ARIMA	2.60	1.7453e-01
DeepDE - CNN	2.20	2.5059e-01
DeepDE - LINREG	6.00	1.7280e-03
DeepDE - LSTM	3.00	1.1719e-01
DeepDE - MLP	3.80	4.7202e-02
DeepDE - NARX	1.60	4.0340e-01
DeepDE - RBF	4.80	1.2186e-02
DeepDE - REGTREE	7.80	4.6328e-05
DeepDE - SVR	1.20	5.3087e-01

Na Figura 7 (a) e (b) é apresentada uma análise comparativa entre os valores reais e previstos, referentes ao conjunto de teste, gerados pelo modelo proposto e pelo melhor preditor em nossa análise, respectivamente. É possível verificar, sub-

jetivamente, um melhor desempenho de previsão do modelo proposto em relação aos melhor modelo investigado neste trabalho.

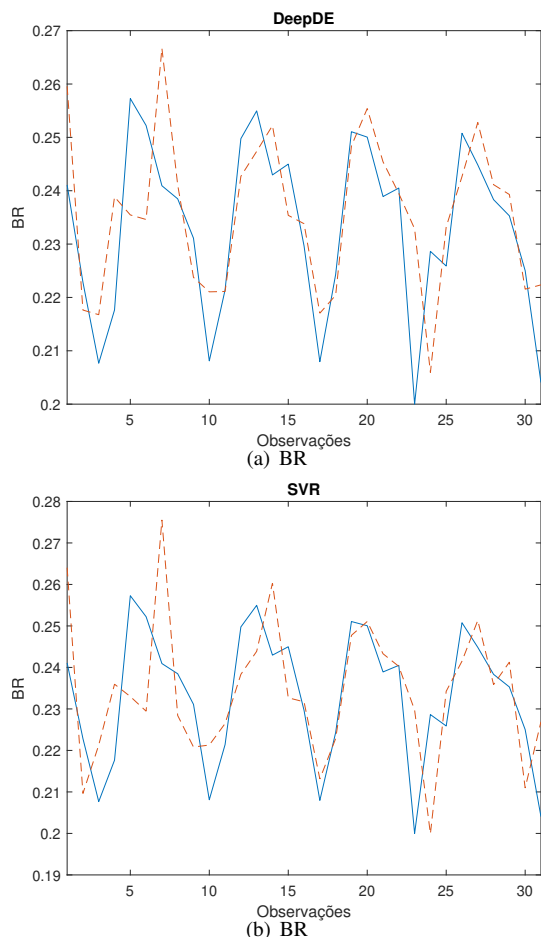


Figura 7. Resultados da previsão do modelo proposto (a) e do melhor predictor investigado (b) para o conjunto de teste.

V. CONCLUSÕES

Neste trabalho, foram apresentadas algumas evidências empíricas de que a série temporal de óbitos por COVID-19 no Brasil é previsível, baseado na análise da função de autocorrelação (para dependência linear) e na informação mútua média (para dependência não-linear) associada ao parâmetro de Hurst (para avaliar a natureza da não-linearidade em termos de componentes de longo ou curto prazo de processos autossimilares).

Com base nessas evidências, é apresentada uma arquitetura de rede neural, denominada *deep dilation-erosion neural network* (DeepDE), disposta em uma estrutura profunda capaz de construir uma modelagem para a série temporal de óbitos por COVID-19 no Brasil. Além disso, para o processo de aprendizado, empregamos um algoritmo de treinamento baseado em gradiente descendente, usando ideias do algoritmo de retropropagação do erro.

Também, é importante mencionar que o modelo proposto foi capaz de capturar padrões relevantes e propriedades variantes no tempo da série temporal investigada. A avaliação de desempenho suportou fortemente a conclusão de que o modelo proposto tem um aprendizado estável e alto poder de generalização para prever a série temporal investigada.

Vale ressaltar que o modelo proposto obteve um desempenho consistentemente superior, considerando a medida MSE, em relação aos outros modelos investigados neste trabalho. Tal fato pode ser atribuído à capacidade do modelo DeepDE em reconstruir eficientemente a dinâmica de óbitos por COVID-19 no Brasil em relação aos outros modelos considerados. Portanto, podemos concluir que o modelo proposto tem uma eficácia muito melhor de maneira geral.

Como trabalho futuro, é importante realizar um estudo sobre a complexidade computacional do modelo proposto e seu processo de aprendizado. Além disso, uma análise experimental adicional deve ser feita com outras séries temporais da COVID-19 para confirmar as evidências encontradas neste trabalho.

REFERÊNCIAS

- [1] G. F. Gao. From “a” to “z”: attacks from emerging and re-emerging pathogens. *Cell*, 172:1157–1159, 2018.
- [2] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, and W. Tan. A novel coronavirus from patients with pneumonia in china, 2019. *N Engl J Med*, 382(2):727–733, 2020.
- [3] C. Lai, T. Shih, W. Ko, H. Tang, and P. Hsueh. Severe acute respiratory syndrome coronavirus 2 (sars-cov-2) and coronavirus disease-2019 (covid-19): The epidemic and the challenges. *International Journal of Antimicrobial Agents*, 55(3):105924, 2020.
- [4] J. She, J. Jiang, L. Ye, L. Hu, C. Bai, and Y. Song. 2019 novel coronavirus of pneumonia in wuhan, china: emerging attack and management strategies. *Clin Transl Med*, 9(1):1–19, 2020.
- [5] F. He, Y. Deng, and W. Li. Coronavirus disease 2019 (covid-19): What we know? *Journal of medical virology*, 92:719–725, 2020.
- [6] J. Cao, X. Hu, W. Cheng, L. Yu, W. Tu, and Q. Liu. Clinical features and short-term outcomes of 18 patients with coronavirus disease 2019 in intensive care unit. *Intensive Care Medicine*, 46:851–853, 2020.
- [7] WHO. Coronavirus disease 2019 (covid-19): Situation report. Technical report, World Health Organization, 2020.
- [8] F. K. Ayittey, M. K. Ayittey, N. B. Chiwero, J. S. Kamasah, and C. Dzuor. Economic impacts of wuhan 2019-ncov on china and the world. *J Med Virol*, 5(92):473–475, 2020.
- [9] Subramanian Rama Iyer and Betty J. Simkins. Covid-19 and the economy: Summary of research and future directions. *Finance Research Letters*, 47:102801, 2022.
- [10] Kathy Leung, Joseph T Wu, and Gabriel M Leung. Effects of adjusting public health, travel, and social measures during the roll-out of covid-19 vaccination: a modelling study. *The Lancet Public Health*, 6(9):e674–e682, 2021.
- [11] S. G. Viveiros Rosa and W. C. Santos. Clinical trials on drug repositioning for covid-19 treatment. *Pan American Journal of Public Health*, 44:1–7, 2020.
- [12] Severino Jefferson Ribeiro da Silva and Lindomar Pena. Collapse of the public health system and the emergence of new variants during the second wave of the covid-19 pandemic in brazil. *One Health*, 13:100287, 2021.

- [13] Daniel R Feikin, Melissa M Higdon, Laith J Abu-Raddad, Nick Andrews, Rafael Araos, Yair Goldberg, Michelle J Groome, Amit Huppert, Katherine L O'Brien, Peter G Smith, Annelies Wilder-Smith, Scott Zeger, Maria Deloria Knoll, and Minal K Patel. Duration of effectiveness of vaccines against sars-cov-2 infection and covid-19 disease: results of a systematic review and meta-regression. *The Lancet*, 399(10328):924–944, 2022.
- [14] Busra Meniz and Mehmet Ozkan. Vaccine selection for covid-19 by ahp and novel vikor hybrid approach with interval type-2 fuzzy sets. *Engineering Applications of Artificial Intelligence*, 119:105812, 2023.
- [15] Amit Kaura, Adam Trickey, Anoop S V Shah, Umberto Benedetto, Ben Glampson, Abdulrahim Mulla, Luca Mercuri, Sanjay Gautama, Ceire E Costelloe, Ian Goodman, Julian Redhead, Kavitha Saravanakumar, Erik Mayer, and Jamil Mayet. Comparing the longer-term effectiveness of a single dose of the pfizer-biontech and oxford-astrazeneca covid-19 vaccines across the age spectrum. *eClinicalMedicine*, 46:101344, 2022.
- [16] Paulo Gabriel Siqueira, Heitor Oliveira Duarte, and Márcio das Chagas Moura. Risk-based cost-benefit analysis of alternative vaccines against covid-19 in brazil: Coronavac vs. astrazeneca vs. pfizer. *Vaccine*, 40(28):3851–3860, 2022.
- [17] Milena Suarez Castillo, Hamid Khaoua, and Noémie Courtejoie. Vaccine effectiveness and duration of protection against symptomatic infections and severe covid-19 outcomes in adults aged 50 years and over, france, january to mid-december 2021. *Global Epidemiology*, 4:100076, 2022.
- [18] Sumit Mohan, Anil Kumar Solanki, Harish Kumar Taluja, Anuradha, and Anuj Singh. Predicting the impact of the third wave of covid-19 in india using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach. *Computers in Biology and Medicine*, 144:105354, 2022.
- [19] Guangheng Wu and Junwei Duan. Blcov: A novel collaborative-competitive broad learning system for covid-19 detection from radiology images. *Engineering Applications of Artificial Intelligence*, 115:105323, 2022.
- [20] Oscar Castillo, Juan R. Castro, Martha Pulido, and Patricia Melin. Interval type-3 fuzzy aggregators for ensembles of neural networks in covid-19 time series prediction. *Engineering Applications of Artificial Intelligence*, 114:105110, 2022.
- [21] T. Ahila and A.C. Subhajini. E-gcs: Detection of covid-19 through classification by attention bottleneck residual network. *Engineering Applications of Artificial Intelligence*, 116:105398, 2022.
- [22] Saima Dil, Nyla Dil, and Zafar H Maken. Covid-19 trends and forecast in the eastern mediterranean region with a particular focus on pakistan. *Cureus*, 12(6), 2020.
- [23] Cheng Zhao, Burcu Tepekule, Nicola G Criscuolo, Pedro D Wendel-Garcia, Matthias P Hilty, Thierry Fumeaux, and Thomas P Van Boeckel. icumonitoring. ch: a platform for short-term forecasting of intensive care unit occupancy during the covid-19 epidemic in switzerland. *Swiss medical weekly*, 150:w20277, 2020.
- [24] Chiara Reno, Jacopo Lenzi, Antonio Navarra, Eleonora Barelli, Davide Gori, Alessandro Lanza, Riccardo Valentini, Biao Tang, and Maria Pia Fantini. Forecasting covid-19-associated hospitalizations under different levels of social distancing in lombardy and emilia-romagna, northern italy: results from an extended seir compartmental model. *Journal of clinical medicine*, 9(5):1492, 2020.
- [25] Duccio Fanelli and Francesco Piazza. Analysis and forecast of covid-19 spreading in china, italy and france. *Chaos, Solitons & Fractals*, 134:109761, 2020.
- [26] B. Tang, N. L. Bragazzi, Q. Li, S. Tang, Y. Xiao, and J. Wu. An updated estimation of the risk of transmission of the novel coronavirus (2019-ncov). *Infect Dis Model*, 5:248–255, 2020.
- [27] B. Tang, X. Wang, Q. Li, N. L. Bragazzi, S. Tang, Y. Xiao, and J. Wu. Estimation of the transmission risk of the 2019-ncov and its implication for public health interventions. *Journal of Clinical Medicine*, 9(2):1–13, 2020.
- [28] J. T. Wu, K. Leung, and G. M. Leung. Nowcasting and forecasting the potential domestic and international spread of the 2019-ncov outbreak originating in wuhan, china: a modelling study. *The Lancet*, 395(10225):689–697, 2020.
- [29] W. C. Roda, M. B. Varughese, D. Han, and M. Y. Li. Why is it difficult to accurately predict the covid-19 epidemic? *Infectious Disease Modelling*, 5:271 – 281, 2020.
- [30] D. Benvenuto, M. Giovanetti, L. Vassallo, S. Angeletti, and M. Ciccozzi. Application of the arima model on the covid-2019 epidemic dataset. *Data in Brief*, 29:105340, 2020.
- [31] Ertugrul Ayyildiz, Melike Erdogan, and Alev Taskin. Forecasting covid-19 recovered cases with artificial neural networks to enable designing an effective blood supply chain. *Computers in Biology and Medicine*, 139:105029, 2021.
- [32] Jayanthi Devaraj, Rajvikram Madurai Elavarasan, Rishi Pugazhendhi, G.M. Shafiuallah, Sumathi Ganesan, Ajay Kaarthic Jeysree, Irfan Ahmad Khan, and Eklas Hossain. Forecasting of covid-19 cases using deep learning models: Is it reliable and practically significant? *Results in Physics*, 21:103817, 2021.
- [33] Weiqiu Jin, Shuqing Dong, Chengqing Yu, and Qingquan Luo. A data-driven hybrid ensemble ai model for covid-19 infection forecast using multiple neural networks and reinforced learning. *Computers in Biology and Medicine*, 146:105560, 2022.
- [34] Manar Abu Talib, Mohamed Abdallah, Abdulrahman Abdeljaber, and Omnia Abu Waraga. Influence of exogenous factors on water demand forecasting models during the covid-19 period. *Engineering Applications of Artificial Intelligence*, 117:105617, 2023.
- [35] Soudeh Ghafouri-Fard, Hossein Mohammad-Rahimi, Parisa Motie, Mohammad A.S. Minabi, Mohammad Taheri, and Saeedeh Nateghinia. Application of machine learning in the prediction of covid-19 daily new cases: A scoping review. *Heliyon*, 7(10):e08143, 2021.
- [36] Muhammad Shoaib, Muhammad Asif Zahoor Raja, Muhammad Touseef Sabir, Ayaz Hussain Bukhari, Hussam Alrabaiah, Zahir Shah, Poom Kumam, and Saeed Islam. A stochastic numerical analysis based on hybrid nar-rbfs networks nonlinear sitr model for novel covid-19 dynamics. *Computer Methods and Programs in Biomedicine*, 202:105973, 2021.
- [37] Abdelkader Dairi, Fouzi Harrou, Abdelhafid Zeroual, Mohamad Mazen Hittawe, and Ying Sun. Comparative study of machine learning methods for covid-19 transmission forecasting. *Journal of Biomedical Informatics*, 118:103791, 2021.
- [38] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel. *Time Series Analysis: Forecasting and Control*. Prentice Hall, New Jersey, third edition, 1994.
- [39] M. B. Stojanovic, M. M. Bozic, M. M. Stankovic, and Z. P. Stajic. A methodology for training set instance selection using mutual information in time series prediction. *Neurocomputing*, 141:236–245, 2014.
- [40] E. Hurst. Long term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116:770–799, 1951.
- [41] R. de A. Araujo, A. L. I. Oliveira, and S. R. de L. Meira. A morphological neural network for binary classification problems. *Engineering Applications of Artificial Intelligence*, 65:12 – 28, 2017.
- [42] S. Haykin. *Neural networks: A comprehensive foundation*. Prentice Hall, New Jersey, 1998.
- [43] Monia Hamdi, Inès Hilali-Jaghdam, Bushra Elamin Elnaim, and Azhari A. Elhag. Forecasting and classification of new cases of covid 19 before vaccination using decision trees and gaussian mixture model. *Alexandria Engineering Journal*, 62:327–333, 2023.
- [44] Yaohao Peng and Mateus Hiro Nagata. An empirical overview of nonlinearity and overfitting in machine learning using covid-19 data. *Chaos, Solitons & Fractals*, 139:110055, 2020.
- [45] M. Friedman. A comparison of alternative tests of significance for the problem of m rankings. *Ann. Math. Statist.*, 11(1):86–92, 03 1940.
- [46] J. W. Tukey. Comparing individual means in the analysis of variance. *Biometrics*, 5:99–114, 1949.