Deep Data Driven Neural Networks for COVID-19 Vaccine Efficacy

Thomas Torku
University Studies Department
Middle Tennessee State University
Murfreesboro, USA
thomas.torku@mtsu.edu

***Abstract*—Vaccination strategies to lessen the impact of the spread of a disease are fundamental to public health authorities and policy makers. The socio-economic benefit of full return to normalcy is the core of such strategies. In this paper, a COVID-19 vaccination model with efficacy rate is developed and analyzed. The epidemiological parameters of the model are learned via a feedforward neural network. An hybrid approach that combines residual neural network with variants of recurrent neural network is implemented and analyzed for reliable and accurate prediction of daily cases. The error metrics and a k-fold cross validation with random splitting reveal that a particular type of the hybrid approach called residual neural network with gated recurrent unit is the best hybrid neural network architecture. The data-driven simulations confirm the fact that the vaccination rate with higher efficacy lowers the infectiousness and basic reproduction number. As a study case, COVID-19 data for the state of Tennessee in USA is used.**

***Index Terms*—Deep Learning,Data-Driven, COVID-19, ResNet, RNN, Vaccination Strategy, k-fold cross validation**

I. INTRODUCTION

The spread of SARS-CoV-2 virus began in Wuhan, China December, 2019 [1]. The virus spread across the globe quickly, affecting many lives and causing deaths, thereby making the World Health Organization (WHO) [1] to declare it a pandemic and naming it COVID-19 [1], [2]. Studies on mitigation measures such as social distancing, quarantining, and government shutdowns for controlling the disease show the need for an effective vaccination program. The rate at which the population is vaccinated and its efficacy are two main factors in an effective vaccination program [3]. There is a need for coordinated efforts by all stakeholders such as policy makers, pharmaceutical companies, researchers, epidemiological and health professionals to effectively champion vaccination campaigns for curbing the spread of the disease [4]. Recent studies on vaccine efficacy [5] and vaccine confidence [6] show the importance of an effective vaccination strategy.

Mathematical models have been used to study the dynamics of an epidemic using systems of differential equations [7]–[9]. The model used in this work is called the SIR model: it divides the population into three groups, namely Susceptible (S), Infected (I), and Recovered (R). The extensions of this basic model have been applied to COVID-19 in recent times [10]. During a pandemic period, the main challenge in this type of model is ascertaining the parameters of the model. Conventional least square method can be used to [11] estimate reasonable constant and time-dependent parameters [12].

Data-driven models that make use of deep learning architectures have been applied to epidemiological models. Raissi et al. [13] used physics informed neural network (PINN) to learn constant parameters of the Influenza model.Kharazmi et al. [12] applied PINN to learn time-dependent parameters of both integer and fractional order epidemiological models. Long et al. [2] is used to learn the time-dependent transmission rate for Susceptible (S), Infected (I), and Recovered (R), and Death (D) (SIRD model), while Grimm et al. [14] presented the effectiveness of using PINN to learn the time-dependent transmission rate of SIR and Susceptible (S), Exposed (E), Infected (I), and Recovered (R) (SEIR) models.

Conventional forecasting techniques such as Auto-Regressive Integrated Moving Average (ARIMA) have been used to make short-term forecasts for confirmed cases of COVID-19 in different countries [15]. Other studies utilized machine learning algorithms, such as Support Vector Machine (SVM), Linear regression, and Exponential Smoothing (ES), to make future predictions of COVID-19 [16]. Recent applications of recurrent neural networks (RNN) with its variant Long Short Term Memory (LSTM) to forecast COVID-19 in Canada, Italy, and USA showed reasonable performance in prediction [17]. The authors in [4] applied LSTM, Bidirectional LSTM (BiLSTM), and Gated Recurrent Unit (GRU) for data-driven simulations of COVID-19 data to different countries. Uncertainties arise from the model, the parameters, and other external factors that impact the accuracy in prediction [11]. There is a need to quantify the uncertainty associated with models [18] for forecasting the trend analysis of COVID-19.

Residual Neural Network (ResNet) is considered, which is capable of learning dynamical systems [19]. The residual connections help to improve accuracy in training and prediction, and are considered as time-stepping techniques modelled after adaptive numerical methods using the concept of integration. Chen et al. [20], in their recent work, showed that the generalized ResNet has the capability of learning intricate unknown dynamic structures of chaotic systems and predicting results with higher accuracy than the standard ResNet structure.

In this paper, a hybrid approach based on residual neural network combined with variants of recurrent neural network for reliable and accurate short-term predictions is considered. The hybrid approach consists of ResNet-LSTM, ResNet-BiLSTM, and ResNet-GRU. The impact of COVID-19 vaccination model with efficacy on each group of population is analyzed. The feedforward neural network solves the inverse problem which incorporates the epidemiological parameters in the loss function. Error metrics for data-driven simulations are discussed to support the claim of effectiveness for hybrid approach. Furthermore, error metrics and a k-fold cross validation with random splitting are discussed. The paper is organized as follows. Section II gives an overview of background in terms of mathematical model including the Susceptible, Infected, Recovered Vaccine model, non-negativity and boundedness of the model, data preprocessing, error metrics and k-fold cross validation. Section III presents the methods used in this work. In section IV, the results are presented and discussion is done in section V.

II. Background

A. Mathematical Model

Given the following population groups

$S(t)$: The individuals that are susceptible per time
$I(t)$: The individuals that are infected per time
$R(t)$: The individuals that are recovered per time
$N(t)$: The total population per time

$$\begin{aligned} \frac{dS(t)}{dt} &= -\beta\frac{S(t)I(t)}{N(t)} - v\eta S(t) \\ \frac{dI(t)}{dt} &= \beta\frac{S(t)I(t)}{N(t)} - \gamma I(t) \\ \frac{dR(t)}{dt} &= \gamma I(t) + v\eta S(t), \end{aligned} \tag{1}$$

with $S(0) \geq 0, I(0) \geq 0, R(0) \geq 0,$ $N(t) = S(t) + I(t) + R(t)$, β transmission rate, γ recovery rate efficacy rate η and vaccination rate v. The population per time is assumed to be constant throughout the vaccination regime.

B. Deep Learning Algorithms

1) Epidemiology Informed Neural Network (EINN): The epidemiology informed neural network (EINN) is inspired by a physics informed neural network (PINN) [13] which incorporates the epidemiological parameters of the model and initial values into the loss function. It is a feedforward neural network that propagates information from the input layer through to the output layer and the loss function incorporates the epidemiological parameters of the model and initial values. The output from EINN satisfies the differential equations in model (1). This is achieved by encoding the the residuals of the model (1) into the loss function. The method of automatic differentiation [21] is used to compute the derivatives of the output with respect to time for each residual equation. In this study, the mean squared error (MSE) is encoded as the loss function which consists of

$$Loss = MSE_{data} + MSE_{sir} + MSE_{U_0}, \tag{2}$$

where

$$\begin{aligned} MSE_{data} =& \frac{1}{M}\sum_{j=1}^{M}||I(t_j) - \hat{I}(t_j)||_2^2 \\ MSE_{sir} =& \frac{1}{M}\sum_{i=1}^{3}\sum_{j=1}^{M}||e_i(t_j,\beta,\gamma)||_2^2 \\ MSE_{U_0} =& \frac{1}{M}\sum_{j=1}^{M}||S_0 - \hat{S}_0||_2^2 \\ &+ \frac{1}{M}\sum_{j=1}^{M}||I_0 - \hat{I}_0||_2^2 \\ &+ \frac{1}{M}\sum_{j=1}^{M}||R_0 - \hat{R}_0||_2^2, \end{aligned} \tag{3}$$

where the residual $e_i, i = 1, \ldots, 3$

$$\begin{aligned} e_1(t_j, \beta, \gamma) &:= \frac{dS(t_j)}{dt_j} + \beta \frac{S(t_j)I(t_j)}{N(t_j)} + \upsilon\eta S(t_j) \\ e_2(t_j, \beta, \gamma) &:= \frac{dI(t_j)}{dt_j} - \beta \frac{S(t_j)I(t_j)}{N(t_j)} + \gamma I(t_j) \\ e_3(t_j, \beta, \gamma) &:= \frac{dR(t_j)}{dt_j} - \gamma I(t_j) - \upsilon\eta S(t_j). \end{aligned} \tag{4}$$

2) Recurrent Neural Network (RNN): Recurrent neural networks are networks that are used to learn problems involving sequential data [4]. In this work, all the variants of RNN namely Long Short Term Memory (LSTM), Bidrectional LSTM (BiLSTM) and Gated Recurrent Unit (GRU) are implemented in PyTorch.

3) Residual Neural Network (ResNet): This type of deep neural network behaves like first order Euler method [19]. In this work we cast the ResNet method as an ODE-solver and construct a hybrid approach with traditional recurrent neural networks. Given a deep neural network of the form

$$y^{out} = \mathbf{N}(y^{in}, \theta).$$

The ResNet is defined as

$$y^{out} = y^{out} + \mathbf{N}(y^{in}\theta).$$

Where y^{out} is the output, y^{in} is the input (data from solver or fabricated or real measurement), $\mathbf{N}$ is the neural network and θ is the parameter of the neural network.

4) Error Metrics: In this section, error metrics for data-driven simulations are presented. A survey of the reason and choice of when to use these error metrics is discussed. The metrics used to measure the performance of regression-based models include: Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and Explained Variance (EV) [4].

III. Methods

The data for this work is obtained from Tennessee Health Department [22]. The available data as of the time of preparing this manuscript is from March 12, 2020 to September 16, 2021. Cumulative cases are used for the training of the EINN where a 7-day rolling mean is applied to the data. In particular, the training data is from March 12, 2020 to December 16, 2020. This is the period of no vaccination program. The data is preprocessed using MinMax scaling. The EINN is implemented in Tensorflow while LSTM, BiLSTM, and GRU as well as k-fold cross validation are implemented in PyTorch.

Figure 1 gives the overall workflow of the deep learning algorithm for prediction. Available COVID-19 data is obtained from reliable sources. It is randomly split into train and test sets of 80% to 20% respectively. The EINN algorithm is used to learn the epidemiology parameters, transmission rate β and recovery rate γ as well as infected group. For reliability and accuracy of the learned parameters, the EINN is run 200 times to quantify the uncertainty in the learned parameters by constructing a confidence interval bound using Bootstrap algorithm. The Poisson distribution is used to resample the data [11]. Using different values of vaccination rates, the EINN learns different β and γ. A standard ODE solver is used to produce epidemic outcomes at different vaccination rates using the learned parameters and efficacy rate η. ResNet data is passed into LSTM, BiLSTM, and GRU to have hybrid method ResNet-LSTM, ResNet-BiLSTM, and ResNet-GRU. The error metrics are used to compute the deviations in the test data. A short term prediction into the future is made for 15 days and a confidence interval is constructed. A cross validation is carried out to evaluate the performance of LSTM, BiLSTM, GRU and the hybrid approach.

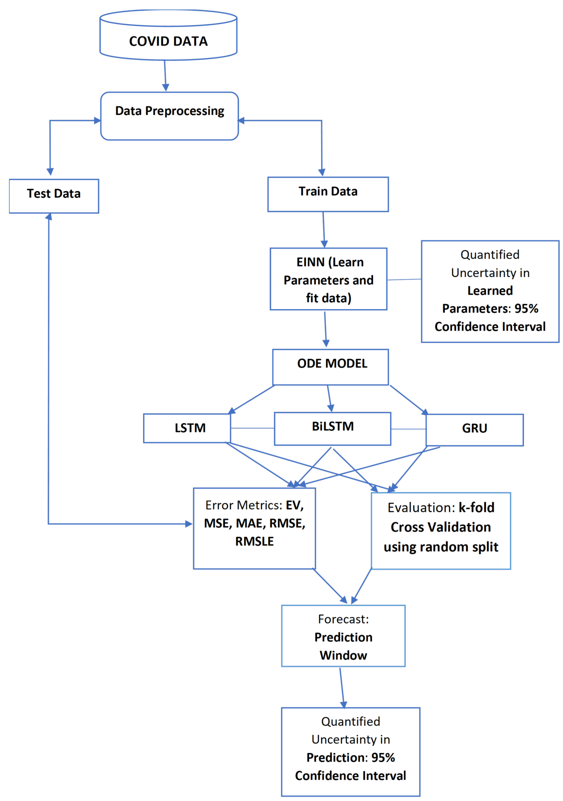

Fig. 1. Deep learning forecasting workflow.

IV. Results

The impact of vaccination with 94% efficacy is presented in Table I. It can be observed from the table that for $v = 0.5\%$, R_0 decreased from 2.5 to 2.0. This means that the higher the efficacy rate, the faster the decline in the spread of the virus. This claim is further supported by the fact that when vaccination rate v increases to 10%, the R_0 decreases to 1.31.

TABLE I
IMPACT OF VACCINATION WITH (94%) EFFICACY RATE.

Vaccine Rate (%)	β	γ	R_0	Infected (%)	Days of Peak
0.00	0.35	0.14	2.5	23.35	46
0.5	0.44	0.22	2.0	15.04	10
1	0.45	0.23	1.96	13.76	9
2	0.42	0.22	1.91	11.92	8
3	0.43	0.24	1.79	10.11	7
6	0.27	0.16	1.89	7.79	5
10	0.17	0.13	1.31	6.57	1

Figure 2 shows results for the hybrid approach: ResNet-LSTM, ResNet-BiLSTM, ResNet-GRU. The outcome ResNet is used as the input data for LSTM, BiLSTM, and GRU. The daily new cases are plotted along with ResNet-LSTM, ResNet-BiLSTM, ResNet-GRU. The outcome from the ResNet architecture is called the ResNet data, which is used to train LSTM, BiLSTM, and GRU. The daily cases and ResNet-LSTM are both colored blue while the ResNet-BiLSTM and ResNet-GRU are colored orange and green respectively. The data spans nine months period: January 16, 2021 to September 16, 2021.

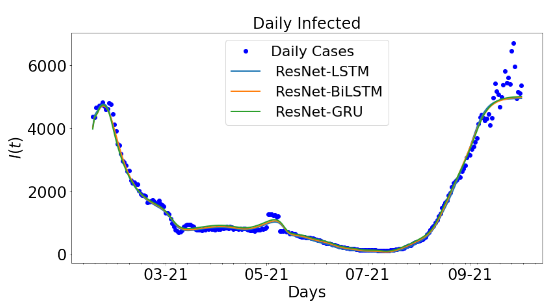

Fig. 2. Data-driven simulation for ResNet-LSTM, ResNet-BiLSTM, and ResNet-GRU.

In Figure 3, the Bootstrap algorithm is used to generate the confidence interval band after 10 independent runs. The lower 95% and upper 95% are given in the graph and the mean prediction is also shown. The Bootstrap algorithm is deployed to run the model 10 times. The Poisson distribution is used to replicate the real COVID-19 data 10 times. For all runs, the mean and standard deviation is computed to ascertain a 95% confidence interval for the future predictions (15 days ahead). The confidence interval bound indicates a reasonable bound for the future predictions.

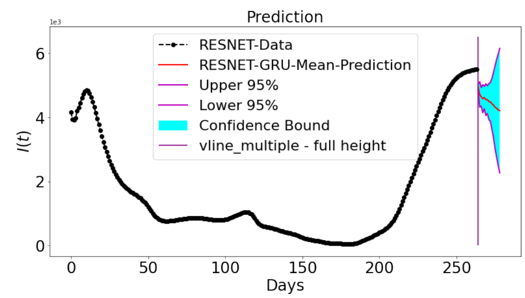

Fig. 3. A confidence interval is constructed for a short-term prediction using ResNet-GRU.

Table II shows the error metrics table for LSTM, BiLSTM, GRU, ResNet-LSTM, ResNet-BiLSTM, and ResNet-GRU. It can be observed that the RMSE values for LSTM, BiLSTM, and GRU are 650.141249, 611.468730, 536.244478 respectively. GRU has the smallest RSME value. Additionally, the MAPE values (relative errors) are 0.159830, 0.150323, 0.131830 respectively for LSTM, BiLSTM, and GRU. The model with the least MAPE value is GRU. Besides, the corresponding EV values for LTSM, BiLSTM, and GRU are 0.891135,0.91171,0.922800. The model with the greatest EV value is GRU. This implies that, using real COVID-19 data, the model with the best error values is GRU. RMSE values for ResNet-LSTM, ResNet-BiLSTM, and ResNet-GRU are 269.863739, 255.473984,238.104349 respectively. The corresponding MAPE and EV values are 0.065988,0.062470,0.058222 and 0.981658,0.981471,0.983856. The hybrid approach with the overall best error values in terms of RMSE, MAPE, and EV is ResNet-GRU.

TABLE II
ERROR METRICS

Approach	RMSE	MAPE	EV
LSTM	650.141 249	0.159 830	0.891 135
BiLSTM	611.468 730	0.150 323	0.911 171
GRU	536.244 478	0.131 830	0.922 800
ResNet-LSTM	269.863 739	0.065 988	0.981 658
ResNet-BiLSTM	255.473 984	0.062 470	0.983 471
ResNet-GRU	238.104 340	0.058 222	0.983 856

Figure 4 demonstrate the mean value of the RMSE. All four values of k are plotted against each value of k for each model in a bar graph for ResNet data. The bar graph shows the cross validation scores for Tennessee. When $k = 4$, the model with lowest RMSE value is ResNet-BiLSTM, followed by ResNet-GRU, then ResNet-LSTM. With $k = 5$, the model with lowest RMSE value is ResNet-GRU, followed by ResNet-BiLSTM, then ResNet-LSTM. When $k = 6$, it can be observed that ResNet-GRU has the lowest RMSE value, followed by ResNet-LSTM and ResNet-BiLSTM. As the value of k increases from 5 to 7, it can be inferred that ResNet-GRU has smallest average RMSE value.

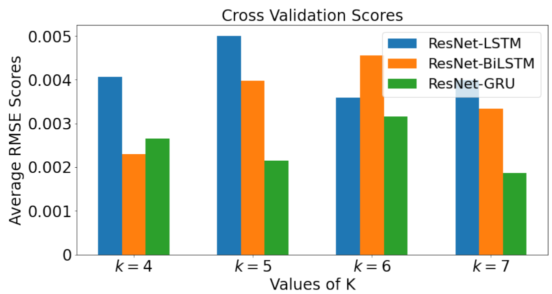

Fig. 4. Cross Validation Scores using mean scores for each value of k for ResNet data.

V. Discussion

In this work, a COVID-19 vaccination model with vaccine efficacy has been developed and analyzed, where the epidemiological parameters of the model were obtained from the inverse problem solved by epidemiological informed neural network. A hybrid approach involving residual neural network with recurrent neural network was implemented. Based on error metric values, the goal of investigating which particular type of the hybrid approach produced reliable future prediction of the infected cases was achieved. In particular, error metric for data-driven simulations demonstrate that ResNet-GRU was the best hybrid approach because of its smallest root mean squared error and mean absolute percentage error values and greatest explained variance value. A k-fold cross validation with random splitting using their average root mean squared error as mean score supported the fact that ResNet-GRU is the best algorithm. Based on quantified uncertainty of 95% confidence interval, the results from short-term prediction demonstrated the overall effectiveness of our approach.

The epidemiological importance of this study was demonstrated in the vaccination model with vaccine efficacy. The model without vaccination revealed a high peak of infectiousness. However, the model with vaccination demonstrated a lower peak of infectiousness. Insights can be drawn from the fact that vaccinating the public using the same vaccination rate per day will curtail the spread of the disease faster. This was confirmed by our study.

VI. Acknowledgements

The author is grateful to the referees for their valuable comments and suggestions which have improved the quality of the paper.

References

[1] S. Ghamizi, R. Rwemalika, M. Cordy, L. Veiber, T. F. Bissyandé, M. Papadakis, J. Klein, and Y. Le Traon, "Data-driven simulation and optimization for COVID-19 exit strategies," in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2020, pp. 3434–3442.

[2] J. Long, A. Khaliq, and K. Furati, "Identification and prediction of time-varying parameters of COVID-19 model: a data-driven deep learning approach," *International Journal of Computer Mathematics*, vol. 98, no. 8, pp. 1617–1632, 2021.

[3] D. Martínez-Rodríguez, G. Gonzalez-Parra, and R.-J. Villanueva, "Analysis of key factors of a SARS-CoV-2 vaccination program: A mathematical modeling approach," *Epidemiologia*, vol. 2, no. 2, pp. 140–161, 2021.

[4] A. Zeroual, F. Harrou, A. Dairi, and Y. Sun, "Deep learning methods for forecasting COVID-19 time-series data: A comparative study," *Chaos, Solitons & Fractals*, vol. 140, pp. 110–121, 2020.

[5] G. Webb, "A COVID-19 epidemic model predicting the effectiveness of vaccination," *Mathematics in Applied Sciences and Engineering*, vol. 2, no. 2, pp. 134–148, 2021.

[6] P. Verger and È. Dubé, "Restoring confidence in vaccines in the COVID-19 era," *Expert Review of Vaccines*, vol. 19, pp. 991 – 993, 2020.

[7] W. O. Kermack and A. G. McKendrick, "A contribution to the mathematical theory of epidemics," *Proceedings of the royal society of London. Series A, Containing papers of a mathematical and physical character*, vol. 115, no. 772, pp. 700–721, 1927.

[8] Y. A. Kuznetsov and C. Piccardi, "Bifurcation analysis of periodic SEIR and SIR epidemic models," *Journal of mathematical biology*, vol. 32, no. 2, pp. 109–121, 1994.

[9] M. H. A. Biswas, L. T. Paiva, and M. De Pinho, "A SEIR model for control of infectious diseases with constraints," *Mathematical Biosciences & Engineering*, vol. 11, no. 4, p. 761, 2014.

[10] Y. Hao, T. Xu, H. Hu, P. Wang, and Y. Bai, "Prediction and analysis of corona virus disease 2019," *PLOS ONE*, vol. 15, no. 10, p. e0239960, 2020.

[11] G. Chowell, "Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts," *Infectious Disease Modelling*, vol. 2, no. 3, pp. 379–398, 2017.

[12] E. Kharazmi, M. Cai, X. Zheng, G. Lin, and G. E. Karniadakis, "Identifiability and predictability of integer-and fractional-order epidemiological models using physics-informed neural networks," *medRxiv*, 2021.

[13] M. Raissi, N. Ramezani, and P. Seshaiyer, "On parameter estimation approaches for predicting disease transmission through optimization, deep learning and statistical inference methods," *Letters in Biomathematics*, vol. 6, no. 2, pp. 1–26, 2019.

[14] V. Grimm, A. Heinlein, A. Klawonn, M. Lanser, and J. Weber, "Estimating the time-dependent contact rate of SIR and SEIR models in mathematical epidemiology using physics-informed neural networks," Universität zu Köln, Technical Report, September 2020. [Online]. Available: https://kups.ub.uni-koeln.de/12159/

[15] T. Dehesh, H. Mardani-Fard, and P. Dehesh, "Forecasting of COVID-19 confirmed cases in different countries with arima models," *MedRxiv*, 2020.

[16] F. Rustam, A. A. Reshi, A. Mehmood, S. Ullah, B.-W. On, W. Aslam, and G. S. Choi, "COVID-19 future forecasting using supervised machine learning models," *IEEE access*, vol. 8, pp. 101 489–101 499, 2020.

[17] V. K. R. Chimmula and L. Zhang, "Time series forecasting of COVID-19 transmission in Canada using LSTM networks," *Chaos, Solitons & Fractals*, vol. 135, p. 109864, 2020.
[18] D. Zhang, L. Lu, L. Guo, and G. E. Karniadakis, "Quantifying total uncertainty in physics-informed neural networks for solving forward and inverse stochastic problems," *Journal of Computational Physics*, vol. 397, p. 108850, 2019.
[19] T. Qin, K. Wu, and D. Xiu, "Data driven governing equations approximation using deep neural networks," *Journal of Computational Physics*, vol. 395, pp. 620–635, 2019.
[20] Z. Chen and D. Xiu, "On generalized residual network for deep learning of unknown dynamical systems," *Journal of Computational Physics*, vol. 438, pp. 110–362, 2021.
[21] L. Lu, X. Meng, Z. Mao, and G. E. Karniadakis, "Deepxde: A deep learning library for solving differential equations," *SIAM Review*, vol. 63, no. 1, pp. 208–228, 2021.
[22] "Tennessee Health Department," https://www.tn.gov/health/cedep/ncov/data, Accessed: 06-28-2021.