Mathematical Biosciences and Engineering

http://www.aimspress.com/journal/mbe

MBE, 20(9): 16807–16823.
DOI: 10.3934/mbe.2023749
Received: 18 July 2023
Revised: 12 August 2023
Accepted: 12 August 2023
Published: 23 August 2023

Research article

Neural-SEIR: A flexible data-driven framework for precise prediction of epidemic disease

Haoyu Wang[1], Xihe Qiu[1,*], Jinghan Yang[1], Qiong Li[2], Xiaoyu Tan[3] and Jingjing Huang[4,5,*]

[1] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China

[2] School of Computer Science and Technology, Beijing Jiaotong University, Beijing 100044, China

[3] INF Technology (Shanghai) Co., Ltd., Shanghai 201203, China

[4] Department of Otolaryngology-Head and Neck Surgery, Eye & ENT Hospital of Fudan University, Shanghai 200031, China

[5] Sleep Disordered Medical Center, Shanghai Municipal Key Clinical Specialty, China

* **Correspondence:** Email: qiuxihe@sues.edu.cn; gennie_xuan@163.com.

Abstract: Accurately modeling and predicting epidemic diseases is crucial to prevent disease transmission and reduce mortality. Due to various unpredictable factors, including population migration, vaccination, control efforts, and seasonal fluctuations, traditional epidemic models that rely on prior knowledge of virus transmission mechanisms may not be sufficient to forecast complex epidemics like coronavirus disease 2019(COVID-19). The application of traditional epidemiological models such as susceptible-exposed-infectious-recovered (SEIR) may face difficulties in accurately predicting such complex epidemics. Data-driven prediction approaches lack the ability to generalize and exhibit low accuracy on small datasets due to their reliance on large amounts of data without incorporating prior knowledge. To overcome this limitation, we introduce a flexible ensemble data-driven framework (Neural-SEIR) that "neuralizes" the SEIR model by approximating the core parameters through neural networks while preserving the propagation structure of SEIR. Neural-SEIR employs long short-term memory (LSTM) neural network to capture complex correlation features, exponential smoothing (ES) to model seasonal information, and prior knowledge from SEIR. By incorporating SEIR parameters into the neural network structure, Neural-SEIR leverages prior knowledge while updating parameters with real-world data. Our experimental results demonstrate that Neural-SEIR outperforms traditional machine learning and epidemiological models, achieving high prediction accuracy and efficiency in forecasting epidemic diseases.

Keywords: Neural-SEIR; long short-term memory; infectious disease prediction; time-series data

1. Introduction

The global morbidity and mortality rates attributable to infectious diseases are 70% more than those attributable to cardiovascular diseases and nearly three times greater than those attributable to cancer, which has a significant impact on society. Accurate forecasting of the epidemic's spread trend is critical for the government to formulate policies, optimize the allocation of medical resources, improve public health safety, and prevent transmission of infectious disease [1].

Conventional epidemiological models have been applied for describing the spread of various infectious diseases. Conventional epidemiological models [2], such as the susceptible infectious susceptible (SIS) [3], the susceptible infectious recovered (SIR) [4], and susceptible exposed infectious recovered (SEIR) models, are effective tools to understand and predict epidemiological trajectories of infectious diseases' outbreaks. However, conventional epidemiological models describe the transmission of infectious disease by modeling the core parameters under an assumed correlation coefficient according to prior knowledge and the forecasting model's parameters are not adaptive, which makes it difficult to model some complex infectious diseases (i.e., COVID-19) due to the complexity and unpredictability of population migration, vaccination, control measures, and seasonal factors [5]. This will result in significant variations from the actual process of forecasting.

Machine learning methods learn from large amounts of data to predict the transmission of epidemics, which have the potential to provide innovative solutions for epidemic prediction. A dozen approaches have been proposed based on the support vector machine(SVM) [6], BP neural network (BPNN) [7], random forest (RF) [8], and convolutional neural network (CNN) [9]. The time-series model(i.e., LSTM), has also demonstrated its superiority in the prediction of epidemics [10], which attracts considerable attention [11].

Despite its success in epidemic forecasting, there are two significant problems with current machine-learning approaches. Most models do not take into account prior knowledge, thus requiring a large amount of data to detect trends in the data in order to obtain accurate predictions; however, it is typically difficult to obtain large amounts of high-quality data, so the aforementioned models have a large prediction bias in practice [12]; furthermore, the aforementioned models do not simultaneously account for time-series information in the data set or do not adequately account for factors, such as seasonality. Overall, current machine learning models directly predict the epidemic using data, without considering disease transmission patterns or prior knowledge, resulting in poor generalizability and low accuracy on small datasets.

To address the limitations of the existing epidemic prediction models, we develop Neural-SEIR, a novel ensemble framework combining a data-driven approach with a traditional SEIR mathematical model. We "neuralize" SEIR and incorporate a priori knowledge of virus transmission into the structure of the model, allowing the model to better utilize this a priori knowledge and account for all relevant practical data. Our method approximates the fundamental parameters using neural networks while preserving the propagation structure of SEIR. The overall framework of the proposed Neural-SEIR is illustrated in Figure 1.

Our major contributions of this paper can be summarized as follows.

- We leverage multiple machine learning methods to overcome the difficulty inherent to the SEIR model, in which many parameters are fixed based on prior knowledge (i.e., infection rate β). After "neuralizing" SEIR, the architecture of the neural process is a data-driven technique that

can update the parameters according to real data and incorporates the prior knowledge to make sequential decisions. Our approach utilizes neural networks to approximate the essential parameters while preserving the SEIR propagation structure, allowing the model to simultaneously leverage a priori knowledge of epidemiological models and capture real-time multidimensional features, hence improving the accuracy of predictions. To the best of our knowledge, we are the first to consider to "neuralize" the inherent structure of SEIR.

- Our Neural-SEIR performs well in inferring uncertainty in the prediction process while also benefiting from past experience, allowing for rapid adjustment of predictions when more data is available. Our model eliminates the requirement for huge amounts of high-quality data for data-driven techniques and performs well with small amounts of real-world data.
- The proposed Neural-SEIR method can substantially reduce the modeling complexity of epidemic simulation and improve the computational efficiency of time-series data. It offers an innovative technique for simulating infectious diseases in complicated environments throughout time.
- We extensively evaluate our proposed method on the public benchmark dataset, and our method outperforms state-of-the-art approaches on epidemic prediction. The proposed ensemble Neural-SEIR can effectively leverage the advantages of time-series models and the prior knowledge of heuristic algorithm to achieve stable and accurate predictions.

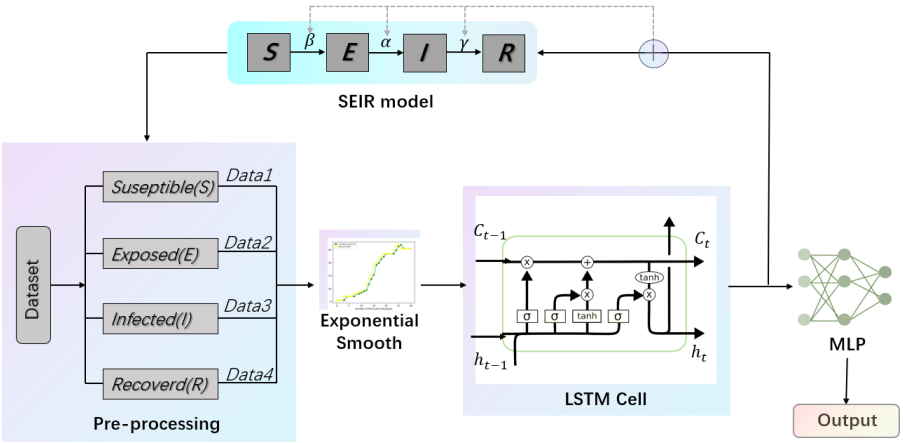

Figure 1. Overall framework of the proposed Neural-SEIR.

2. Related work

2.1. Traditional epidemiological models

The SIS model is widely used for describing the transmission dynamics of infectious diseases [13]. It continues to serve as the foundation for numerous infectious disease models, and is applied in epidemic prevention and control decision-making [3]. The SIR model [14] is an extension of the SIS model that considers the possibility of recovery from infection. It has been extensively applied to simulate smallpox, hepatitis, measles [15], and other immunogenic infectious diseases [16] in predicting disease spread and developing effective strategies for managing disease outbreaks. SEIR model [17] is an extension of the SIR model, considering susceptible, exposed, infectious, and

recovered individuals [18] and the incubation period. It takes the effects of various factors [19] of vaccinations, government control measures [20], and mutations during an epidemic [21] into account. Epidemiological models employ mathematical modeling, numerical simulations, data validation, and statistical techniques [22] to account for diverse aspects of an epidemic. Nevertheless, these models are subject to certain limitations, such as assuming a single maximum peak and fixed transmission rates [23]. To accurately capture the complex dynamics of an epidemic, it is crucial to incorporate real-world factors, including socioeconomic variables [24], epidemic control measures, and vaccination strategies, into an ideal predictive model [25].

2.2. Exponential smoothing methods and applications

Exponential smoothing [26] is a weighted moving average technique that assigns varying weights to observations at different times [27], emphasizing recent observations and enabling faster reflection of market changes in forecasts [28]. This scalable method can modify the rate of weight change, adjusting the degree of smoothing for trend changes and improving the homogeneity of time series observations [29]. Exponential smoothing is widely applied in epidemiological models, such as the forecasting of quarterly incidence of black fever [30], the hepatitis inflammation [31, 32], hand, foot and mouth disease [33], which benefits from the model's ability to estimate cyclical trends through its incorporation of linear and seasonal trends.

2.3. Time-series forecasting models

Sequential models are suitable for predicting the dynamic nature of epidemic data such as COVID-19, which exhibits typical time-series features that require the incorporation of historical information over the incubation period. Incorporating updating data with historical knowledge can improve the accuracy of predictions [34]. LSTM [35] is a powerful method for processing sequence data, capable of recognizing patterns in historical or future data. Its ability to manage complex time-series dependencies has enabled successful applications across various domains, including speech recognition [36], weather prediction [37], energy prediction [38], financial trend forecast [39], drug prediction [40], and language modeling. Dutta *et al.* [41] combine RNN and LSTM to forecast COVID-19 cases. Reddy *et al.* [10] employ an LSTM model to forecast the transmission of COVID-19 in Canada, with high accuracy of 93.4% (short-term) and 92.2% (long-term). Shwet *et al.* [11] present a CNN-LSTM hybrid deep learning prediction model that accurately forecasts the COVID-19 epidemic in India, demonstrating that adding additional convolutional layers to the LSTM layer may enhance the performance of the forecasting model. The data-driven nature of LSTM models necessitates a substantial amount of data to achieve accurate predictions, making them prone to errors when applied for generalization purposes. However, it is important to note that crucial factors such as population density, medical treatment level, population movement, quarantine rates, and virus mutation rates are often not adequately considered in these models. These factors play a significant role in shaping the dynamics of infectious diseases and can greatly impact the accuracy of predictions. Therefore, it is essential to incorporate these factors into the modeling framework to enhance reliability and applicability in infectious disease prediction.

3. Methods

Exponential smoothing structure is a technique used to improve data observation by adjusting the weights' rate of change. It involves utilizing various values of a to forecast values and modify the weights accordingly. a is a parameter that determines the weight assigned to each observation in the exponential smoothing process.

We first construct an exponential smoothing structure to improve the observation of the data by utilizing various values of a for the forecast values, thereby modifying the weights' rate of change. Exponential smoothing enables the data to be scaled to the assigned weights, thereby increasing the weight of the observations and enabling the predicted values to rapidly reflect the actual changes in the epidemic data. We add it to each category of the already pre-process dataset to facilitate subsequent model learning. Figure 2 demonstrate the effects of smoothing on the Susceptible and Removed data, respectively.

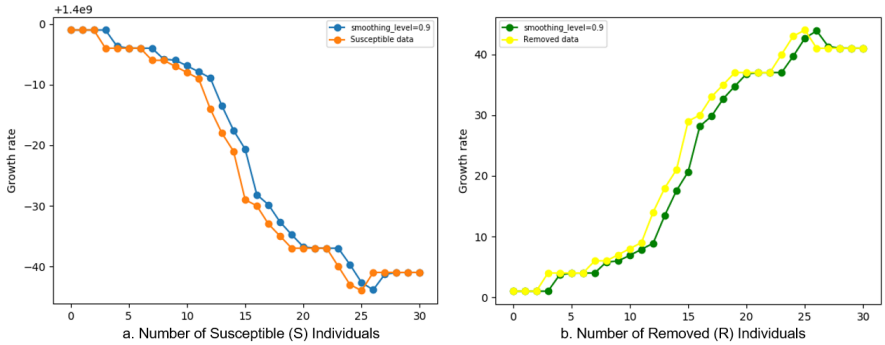

Figure 2. (a)The number of Susceptible people after exponentially smoothing. The orange dots indicate the original data, while the blue dot indicate the effects of exponentially smoothing level at 0.9. (b)The number of Removed individuals after exponentially smoothing. The orange dots indicate the original data, while the blue dot indicate the effects of exponentially smoothing level at 0.9.

S7, I7, and E7 represents the variables corresponding to the seven most recent days of the forecast. SEIR is an acronym for the compartmental model used in epidemiology, representing the variables for Susceptible (S), Exposed (E), Infected (I), and Removed (R) individuals.

By exponentially smoothing S_7, I_7, and E_7, we obtain the seven most recent days of the forecast (which can be any number of days). The final iteration of SEIR is used to determine the forecast values for the last seven days, with the original iteration equation being described as follows.

$$E_n = E_{n-1} + r\beta I_{n-1}S_{n-1}/N - \alpha E_{n-1} + r_2\beta_2 E_{n-1}S_{n-1}/N \tag{3.1}$$

$$I_n = I_{n-1} + \alpha E_{n-1} - \gamma I_{n-1} \tag{3.2}$$

$$R_n = R_{n-1} + \gamma I_{n-1} \tag{3.3}$$

The parameters of the infectious disease model are defined as follows: β represents the infection rate and denotes the probability of disease transmission between susceptible and infectious populations; α

denotes the incubation rate, which is the rate of exposed individuals becoming infected; γ denotes the recovery rate and is determined by the average duration of an infection.

To capture the characteristics of infectious disease transmission, the neural networks take into account various factors and variables that influence the spread of the disease, such as age distribution, population density, temperature, air quality, and vaccination campaigns. These factors play a crucial role in determining the susceptibility, transmission of the infectious agent, and the overall dynamics of the epidemic.

The neural networks approximate these core parameters by learning the complex relationships between the input variables and the observed epidemic patterns. Through training and optimization, the neural networks aim to capture the underlying dynamics of disease transmission and provide accurate predictions of future epidemic trends.

In order to predict the value of S, we substitute $r\beta$ with $\exp_1(NN(x))$ and $r_2\beta_2$ with $\exp_2(NN(x))$, where $\exp_1(NN(x))$ is the value predicted using the LSTM. The variable x in $\exp_1(NN(x))$ corresponds to the confirmed number of I, while in $\exp_2(NN(x))$ it corresponds to the confirmed number of S. Consequently, we derive the following formula.

$$S_n = S_{n-1} - \exp_1(NN(x))li_{n-1}ls_{n-1}/N - \exp_s(NN(x))le_{n-1}ls_{n-1}/N \tag{3.4}$$

In accordance with equation 3.2, I_7 is derived from li_{n-1} and S_7 is derived from ls_{n-1}, which are both corrected and updated based on historical data.

Algorithm 1 Total Algorithm

Input: Susceptible S, Exposed E, Infected I, Removed R, number of representative instances n, batch size N, Variable parameters γ and β.

Output: Predicted Infected value labels l_t, Forecast values for the next few days $\hat{y}_{t+1.t+h}$.

The dataset containing the four types of individual S, E, I, and R labels are subjected to data smoothing, thereby facilitating the subsequent model's learning process;

Plugging the processed data into the LSTM time series model for feature extraction and learning;

$C_t = C_{t-1} * \sigma(f_1(h_{t-1}, x_t))$;

$h_t = \tanh(C_t) * \sigma(f_4(h_{t-1}, x_t))$;

Update C_t and h_t by training in Memory gates and forgetting gates in LSTM;

The higher dimensional feature set containing the time-series features is fed into our Neural-SEIR for learning and the final result is obtained;

Train the model using α and γ;

Repeat: Query instance-label to update S, E, I and R;

Compute l_t and $\hat{y}_{t+1.t+h}$ using the prediction network;

Calculating model error assessment values to f;

Update the prediction network using α and γ;

Until ($|\mathbf{f}| \leq f$)

The combination of exponential smoothing and LSTM involves using exponential smoothing to preprocess the input data before feeding it into the LSTM network. This preprocessing step helps to remove noise and capture the overall trend and seasonality in the data. The preprocessed data is then

used as input to the LSTM network, which learns to model the complex temporal dependencies and make predictions based on the historical information.

The LSTM architecture addresses the "memory" limitations of the vanilla RNN model by using a neural network composed of node cells and changing features that represent the learning network layers, connected by a linear merge representation, allowing the transmitted feature learning content to be replicated at multiple locations. The LSTM model processes the input data S and I by first passing them through the forgetting gate, which determines how much of the information from the previous memory cell should be forgotten. Input gates control which of the currently calculated S and I can be updated to the memory cell in the form of a new state.

$$C_t = C_{t-1} + \sigma\left(f\left(h_{t-1}, x_t\right)\right) * \tanh\left(f\left(h_{t-1}, x_t\right)\right) \tag{3.5}$$

Where C_t represents the current value at time t, C_{t-1} represents the previous value at time $t-1$. σ denotes the sigmoid function, which squashes the input values between 0 and 1. $f(h_{t-1}, x_t)$ represents a function that takes as input the previous hidden state h_{t-1} and the current input x_t. The equation states that the current value C_t is obtained by adding the previous value C_{t-1} with the product of $\sigma\left(f\left(h_{t-1}, x_t\right)\right)$ and $\tanh\left(f\left(h_{t-1}, x_t\right)\right)$.

Through the LSTM's memory gate, the important knowledge will be recorded while forgetting irrelevant knowledge, the important knowledge learned by LSTM will be transmitted to the subsequent SEIR structure.

$$h_t = \tanh\left(C_t\right) * \sigma\left(f\left(h_{t-1}, x_t\right)\right) \tag{3.6}$$

The LSTM neural network, equipped with three gates, can modify cellular states by adding or removing information. The limitations of conventional epidemiological epidemic prediction models are attributable to their failure to account for the dynamic nature of data. LSTM has some degree of functionality similar to GRU. Although LSTM and GRU differ in their internal structures, they are designed to address the same problem and demonstrate similar characteristics when dealing with sequential data. This similarity in functionality allows LSTM to enhance its performance in capturing the dynamic nature of epidemiological data.

Overall, the LSTM's capability to modify cellular states, coupled with its ability to handle long-term dependencies and capture sequential patterns, makes it a valuable tool for improving the accuracy and effectiveness of epidemic prediction models.

$$r_t = \sigma\left(f\left(h_{t-1}, x_t\right)\right) \tag{3.7}$$

$$h_t^c = \tanh\left(f\left(r_t * h_{t-1}, x_t\right)\right) \tag{3.8}$$

Where r_t serves as a gate control variable that regulates the flow of information. Its value ranges between 0 and 1, determining the extent to which the previous hidden state h_{t-1} and the current input x_t influence the current hidden state h_t^c. h_t^c represents the current hidden state. It is computed based on the gate variable r_t, the previous hidden state h_{t-1}, and the current input x_t.

We utilize the sigmoid and tanh functions as activation functions in our neural network architecture. The tanh function is applied to generate candidate memories, and any superfluous data is normalized to a range of -1 to 1 using this function before being updated by the output of the sigmoid function. The sigmoid function determines which current and past input data values should be updated to compute

the cell state data. The tanh layer produces a candidate vector which is then added to the cell state. These two vectors are combined to produce updated values that are further processed to generate the final output. The sigmoid function determines the output cell states. In addition, the tanh function has a larger gradient around an input of 0 than the sigmoid function, leading to faster convergence.

To account for uncertainty, the Neural-SEIR framework incorporates probabilistic modeling approaches. This involves considering probability distributions for the input variables or parameters instead of fixed values. By sampling from these distributions, multiple simulations can be performed, resulting in a range of possible epidemic outcomes.

Parameter estimation uncertainty refers to the uncertainty in estimating the values of model parameters from the available data. We use maximum likelihood estimation which provides estimates of the parameters along with uncertainty measures.

The optimal result is obtained when $\alpha = 0.13$ and $\gamma = 0.066$. Finally, we predict the values of S, E, I and R for the next 7 days.

$$l_t = \alpha y_t + (1 - \alpha) l_{t-1} \tag{3.9}$$

$$\hat{y}_{t+1.t+h} = \exp(NN(x)) * l_t \tag{3.10}$$

Where l_t represents the predicted number of infection individuals at time t, which is determined by the previous step's l_{t-1} and weight coefficients α. The predicted values $\hat{y}t + 1$ and $\hat{y}t + h$ are derived based on the present value of l_t, indicating a correlation between the predicted values and l_t.

The overview of the "neuralizing" process is illustrated in Figure 3. The total algorithm is described in Algorithm 1.

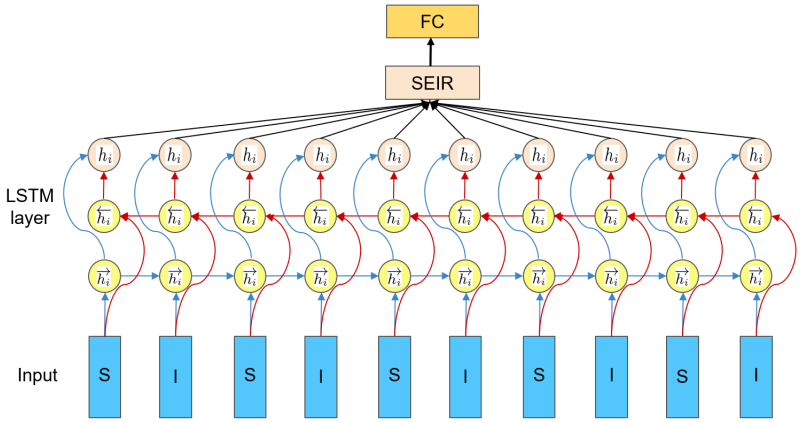

Figure 3. The"neuralizing"process, the input data have been pre-processed by exponential smoothing.

4. Experimental results and analysis

In this section, the dataset, the experimental details, results, and analysis of the proposed model are demonstrated. The experiments are conducted on NVIDIA GeForce RTX 3090 hardware and encoded with Python Tensorflow 1.8.1.

4.1. Data description

For this study, we have utilized the Covid-19 dataset that contains information on the confirmed cases, deaths, and recovered cases in Guangdong Province, China*. We select cases from December 2019 to February 2020. Before conducting the experiments, we preprocessed the data, including filing and deduplication. We also categorize the data situations into four categories, S, E, I, and R, respectively. In the SEIR model, each symbol represents a specific group of individuals within a population during the course of an infectious disease outbreak.

- Susceptible (S): This group includes individuals who are susceptible to the disease and can become infected if exposed to the pathogen. They have not yet been infected and do not have any immunity.
- Exposed (E): This group consists of individuals who have been exposed to the pathogen but are not yet infectious. They have been infected but are in the latent or incubation period, during which the virus is multiplying within their bodies, but they are not yet showing symptoms or capable of transmitting the disease.
- Infectious (I): This group comprises individuals who have been infected, are capable of transmitting the disease, and are showing symptoms. They can spread the pathogen to susceptible individuals through direct or indirect contact.
- Recovered (R): This group includes individuals who have recovered from the infection and have developed immunity to the disease. They are no longer susceptible to reinfection and cannot transmit the pathogen to others.

The SEIR model describes the process of disease transmission by tracking the flow of individuals between these four groups over time. Initially, the disease spreads from the susceptible population to the exposed population through contact with infectious individuals. The exposed individuals then progress to the infectious stage, where they can transmit the disease to susceptible individuals. Over time, some infected individuals recover and move into the recovered group, while others may unfortunately succumb to the disease.

By incorporating the dynamics of these four groups, the SEIR model provides insights into the progression and control of infectious diseases, helping researchers and policymakers understand the impact of interventions such as vaccination, social distancing, and quarantine measures.

4.2. Main results

Accurate prediction of epidemics is a complex task that involves various factors and uncertainties. While many researchers have made significant efforts in this area, it is crucial to acknowledge the challenges and limitations that contribute to the less-than-optimal performance of existing prediction models.

One reason for the suboptimal performance of epidemic prediction models is the inherent complexity of infectious diseases. Epidemics are influenced by numerous dynamic factors, including the characteristics of the pathogen, the behavior of the host population, and the effectiveness of control measures. These factors interact in intricate ways, making it difficult to capture the full complexity of the system and accurately predict its future behavior.

*The data used in this study can be found at http://www.dxy.cn/.

Additionally, data availability and quality pose significant challenges. Accurate prediction models rely on comprehensive and reliable data, including information on disease prevalence, transmission rates, and population demographics. However, data collection and reporting systems may be incomplete, inconsistent, or delayed, leading to uncertainties and inaccuracies in the input data for prediction models.

Furthermore, the evolving nature of epidemics presents a challenge. Outbreaks can exhibit sudden changes in transmission patterns, affected populations, or the emergence of new variants. These dynamic shifts can render existing prediction models less effective, as they may not account for the rapid changes in the epidemic landscape.

Based on the prior knowledge of the SEIR model, we classify the results into four categories: S, E, I, and R. Figure 4 displays Neural-SEIR model's prediction results for the number of Susceptible(S), Exposed(E), Infected(I), and Removed(R) individuals, respectively. The results indicate that our Neural-SEIR model accurately captures the dynamics of the COVID-19 epidemic. Specifically, infected individuals increase exponentially during virus growth, peak after several weeks, reaching an inflection point, and gradually decrease until elimination, with concurrent stabilization of E and R numbers. Figure 5 exhibits the superior predictive performance of the proposed Neural-SEIR model

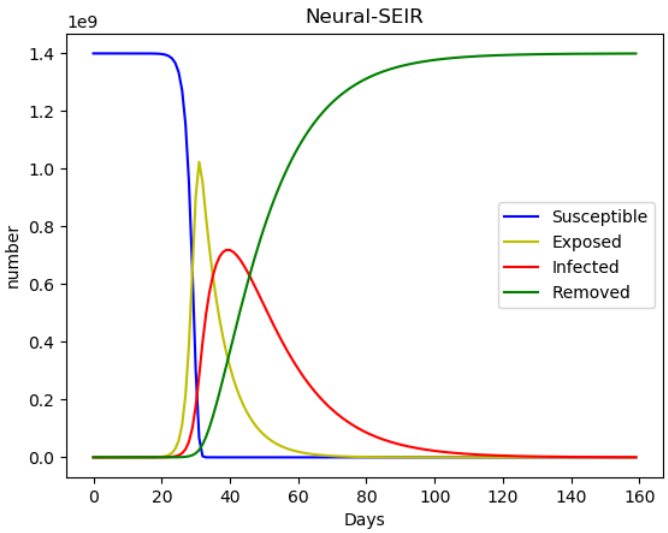

Figure 4. The predictions for the number of S, E, I, and R individuals using Neural-SEIR.

in forecasting the number of infected individuals. The model's effectiveness in predicting future trends in infectious diseases is supported by the small deviations between predicted and actual values. The predictive accuracy is superior to two widely used benchmarks(i.e., LSTM and DeepAR) time-series sequential prediction models. The LSTM model overestimates the number of infected individuals, resulting in a significant discrepancy between predicted and actual values over time. Conversely, the DeepAR model underestimates the number of infected individuals, leading to a lower number of predicted cases than the actual cases. Our findings suggest that the Neural-SEIR model is a promising tool for predicting the spread of infectious diseases with a high level of accuracy.

We evaluate the performance of our proposed model utilizing root mean square error(RMSE), symmetric mean absolute percentage error(SMAPE) and mean absolute error(MAE) evaluation

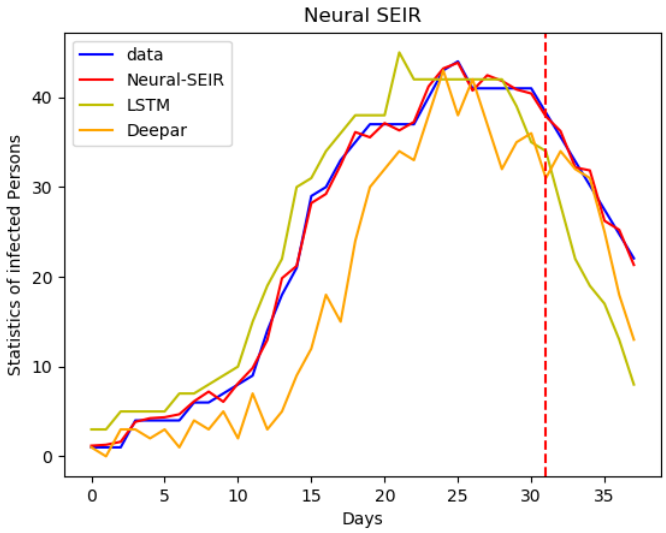

Figure 5. Prediction performance of infected(I) individuals among various models.

indicators. RMSE measures the deviation between an observation and the true value, which is defined as:

$$\mathrm{RMSE}(\mathrm{X}, \mathrm{h}) = \sqrt{\frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} \left(\mathrm{h}\left(\mathrm{x}^{(\mathrm{i})}\right) - \mathrm{y}^{(\mathrm{i})}\right)^2} \tag{4.1}$$

SMAPE describes the accuracy of the time-series models, which is defined as:

$$SMAPE = \frac{100\%}{n} \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{(|\hat{y}_i| + |y_i|)/2} \tag{4.2}$$

MAE is the average of the absolute errors, which describes the prediction error, which is defined as:

$$\mathrm{MAE}(\mathrm{X}, \mathrm{h}) = \frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} \left|\mathrm{h}\left(\mathrm{x}^{(\mathrm{i})}\right) - \mathrm{y}^{(\mathrm{i})}\right| \tag{4.3}$$

In Table 1, our model is compared to time-series models (i.e., LSTM and DeepAR) and the SEIR epidemiological model, with all indicators showing our model's superiority over the benchmarks. These findings provide significant evidence of the efficiency of our proposed model in forecasting the number of infected individuals, which is the most important factor in epidemic transmission. The results suggest that the proposed model can be utilized as a reliable tool for forecasting the spread of infectious diseases in different settings.

4.3. Ablation study

We conduct extensive ablation studies to investigate the impact of each component of our proposed method.

Table 1. Predictive performance results of different methods.

	rmse	smape	mae
LSTM	7.98	**18.48**	7.18
SEIR	49.02	198.96	45.79
DeepAR	24.87	96.92	23.34
Neural-SEIR	**6.49**	19.13	**5.96**

4.3.1. The Impact of Exponential Smoothing Parameters

We investigate the effects of various exponential smoothing parameters on the COVID-19 dataset. Table 2 displays the effects of different exponential smoothing parameters on the COVID-19 dataset, with the optimal smoothing level found to be 0.9. The smoothing index method is essentially a special type of weighted average method. Exponential smoothing assigns greater weight to data points that are closer together when predicting future observations using a weighted average. The smoothing index a determines the level of influence that actual data has on the prediction results, with higher values having a greater impact and lower values having a lesser impact on the forecast results. Hence, a greater smoothing exponent represents a short-term trend in data values, whereas a smaller smoothing exponent reflects a long-term trend of constant data values. Thus, a large smoothing index is optimal for predicting COVID-19 precisely in the short term, where data values are highly fluctuating.

Table 2. Predictive performance results of smoothing level.

	rmse	smape	mae
0.3	28.69	37.31	19.88
0.6	15.13	24.85	13.42
0.9	**6.49**	**19.12**	**5.96**
1.2	8.17	20.52	6.34

4.3.2. The Predictive Performance in Forecasting Different Days

In public decision-making for epidemic transmission prevention and control, such as in the case of COVID-19, it is often crucial to analyze the forecast results for periods of 1, 3, and 7 days. This analysis is essential for allocating medical resources between regions and adjusting prevention strategies (i.e., isolation measures). In this regard, we perform ablation experiments to evaluate the predictive performance of our proposed model across different forecast periods.

We present the results of our experiments in Table 3, which demonstrate the effectiveness of our proposed model in short-term forecasting, particularly in predicting the next 1 day. The ability to accurately forecast in the short-term is particularly valuable in emergency situations where medical resource allocation is critical. Our findings suggest that our model can provide reliable and timely forecasts to inform decision-making in the management of infectious diseases like COVID-19.

4.3.3. The Impact of Different Learning Model

We analyze the key components of our model in Table 4. As shown in the table, our proposed Neural-SEIR model outperforms ES with LSTM and LSTM with SEIR model in terms of all evaluation

Table 3. Predictive performance results of different days.

	rmse	smape	mae
1 day	**2.71**	**6.83**	**2.71**
3 day	4.58	11.79	4.41
7 day	6.49	19.12	5.97

indicators. The superior effects can be attributed to the ensembling of SEIR, ES, and LSTM. The SEIR model can obtain a priori knowledge, the ES model exhibits superiority in capturing seasonal information, and the LSTM to acquire other time series features.

Table 4. Predictive performance results of our model.

	rmse	smape	mae
ES+LSTM	9.56	21.09	8.38
LSTM+SEIR	15.72	31.56	13.64
Neural-SEIR	**6.49**	**19.12**	**5.97**

5. Conclusion

In this paper, we present a novel ensemble data-driven forecasting framework, Neural-SEIR, to achieve precise prediction of infectious diseases. Through the "neuralizing" SEIR mathematical model, we incorporate the real-time practical information into the propagation structure, thus allowing the model to leverage both a priori knowledge and practical information, with high precision and learning efficiency. We extensively conduct experiments to evaluate our proposed method on the real-world dataset. Our proposed method outperforms other state-of-the-art time series models in predicting disease types, with a lower RMSE (6.49) and MAE (5.97). Overall, the experimental results indicate that our approach has the potential to be used to better assist physicians in early intervention and clinical prevention of COVID-19, which might help reduce the financial and resource costs of the public health system. In the future, we will incorporate additional variables, and diverse data sources can improve its ability to capture the complex dynamics of disease transmission and make more accurate predictions. Furthermore, integrating the Neural-SEIR framework into real-time monitoring systems and decision support tools will provide up-to-date epidemic predictions for informed decision-making and timely interventions to improve disease control strategies.

Use of AI tools declaration

The authors declare they have not used Artificial Intelligence (AI) tools in the creation of this article.

Acknowledgments

This work is supported by the National Natural Science Foundation of China, "Science and Technology Innovation Action Plan" Shanghai Natural Science Foundation, Shanghai Municipal Key

Clinical Specialty, China, Eye & ENT Hospital's double priority project A (Grant No.62102241, No.23ZR1425400, No.shslczdzk00801, SYA202210 to Dr. Huang and Dr. Wei).

Data availability

The datasets generated during and analyzed during the current study are available from the corresponding author upon reasonable request.

Conflict of interest

The authors declare there is no conflict of interest.

References

1. R. Sabino-Silva, A.C.G. Jardim, W.L. Siqueira, Coronavirus covid-19 impacts to dentistry and potential salivary diagnosis, *Clin. Oral. Invest*, **24** (2020), 1619–1621. https://doi.org/10.1007/s00784-020-03248-x
2. X. Wang, Z. Wang, H. Shen, Dynamical analysis of a discrete-time sis epidemic model on complex networks, *Appl. Math. Lett*, **94** (2019), 292–299. https://doi.org/10.1016/j.aml.2019.03.011
3. X. Meng, S. Zhao, T. Feng, T. Zhang, Dynamics of a novel nonlinear stochastic sis epidemic model with double epidemic hypothesis, *J. Math. Anal. Appl.*, **7** (2015), 227–242. https://doi.org/10.1016/j.jmaa.2015.07.056
4. N. Sene, Sir epidemic model with mittag–leffler fractional derivative, *Chaos. Soliton. Fract.*, **137** (2020). https://doi.org/10.1016/j.chaos.2020.109833
5. D. Courtney, P. Watson, M. Battaglia, B. H. Mulsant, P. Szatmari, Covid-19 impacts on child and youth anxiety and depression: Challenges and opportunities, *Can. J. Psych.*, **10** (2020). https://doi.org/10.1177/0706743720935646
6. P. Khanna, S. Kumar, Malaria parasite classification employing chan–vese algorithm and svm for healthcare, *IC4S*, (2019), 697-711. https://doi.org/10.1007/978-981-15-3369-3_51
7. J. W. Tian, Y. Liu, W. F. Zheng, L. R. Yin, Smog prediction based on the deep belief - BP neural network model (DBN-BP), *Urban. Clim.*, **41** (2022). https://doi.org/10.1016/j.uclim.2021.101078
8. M. O. Edeh, S. Dalal, I. C.Obagbuwa, B. V. V. S. Prasad, S. Z. Ninoria, M. A. Wajid, et al., Bootstrapping random forest and chaid for prediction of white spot disease among shrimp farmers, *SCI. Rep.-UK*. https://doi.org/10.1038/s41598-022-25109-1
9. C. J. Huang, Y. H. Chen, Y. Ma, P. H. Kuo, Multiple-input deep convolutional neural network model for covid-19 forecasting in china, *MedRxiv* (2020). https://doi.org/10.1101/2020.03.23.20041608
10. V. K. R. Chimmula, L. Zhang, Time series forecasting of covid-19 transmission in canada using lstm networks, *Chaos. Soliton. Fract.*, **135** (2020), 109864. https://doi.org/10.1016/j.chaos.2020.109864

11. S. Ketu, P. K. Mishra, India perspective: Cnn-lstm hybrid deep learning model-based covid-19 prediction and current status of medical resource availability, *Soft. Comput.*, **26** (2022), 645—664. https://doi.org/10.1007/s00500-021-06490-x

12. Q. Ni, J. Kang, M. Tang, Y. Liu, Y. Zou, Learning epidemic threshold in complex networks by convolutional neural network, *Chaos*, **29** (2019), 113106. https://doi.org/10.1063/1.5121401

13. S. Jafarizadeh, D. Veitch, Optimal curing resource allocation for epidemic spreading processes. *Automatica*, **150** (2023), 110851. https://doi.org/10.1016/j.automatica.2023.110851

14. R. Engbert, M. M. Rabe, R. Kliegl, S. Reich, Sequential data assimilation of the stochastic seir epidemic model for regional covid-19 dynamics, *B. Math. Biol.*, **83** (2021). https://doi.org/10.1007/s11538-020-00834-8

15. N. S. Barlow, S. J. Weinstein, Corrigendum to "accurate closed-form solution of the sir epidemic model" [physica d 408 (2020) 132540], *PHYSICA. D*, **416** (2020), 132807. https://doi.org/10.1016/j.physd.2020.132540

16. K. M. A. Kabir, K. Kuga, J. Tanimoto, Analysis of sir epidemic model with information spreading of awareness, *Chaos. Soliton. Fract.*, **119** (2019), 118-–125. https://doi.org/10.1016/j.chaos.2018.12.017

17. A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, et al., Early dynamics of transmission and control of covid-19: a mathematical modelling study, *Lancet. Infect. Dis.*, **20** (2020), 553—558. https://doi.org/10.1016/S1473-3099(20)30144-4

18. K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, et al., The effect of control strategies to reduce social mixing on outcomes of the covid-19 epidemic in wuhan, china: A modelling study, *Lancet. Public. Health*, **5** (2020), 261—270. https://doi.org/10.1016/S2468-2667(20)30073-6

19. J. L. Sainz-Pardo, J. Valero, Covid-19 and other viruses: Holding back its spreading by massive testing, *Expert. Syst. Appl.*, **186** (2021), 115710. https://doi.org/10.1016/j.eswa.2021.115710

20. T. Phan, S. Brozak, B. Pell, A. Gitter, A. Xiao, K. D. Menad, et al., A simple SEIR-V model to estimate COVID-19 prevalence and predict SARS-CoV-2 transmission using wastewater-based surveillance data, *Sci. Total. Environ.*, (2022). https://doi.org/10.1016/j.scitotenv.2022.159326

21. P. Jithesh, A model based on cellular automata for investigating the impact of lockdown, migration and vaccination on covid-19 dynamics, *Comput. Meth. Prog. Biol.*, **211** (2021), 106402. https://doi.org/10.1016/j.cmpb.2021.106402

22. L. López, X. Rodo, A modified seir model to predict the covid-19 outbreak in spain and italy: Simulating control scenarios and multi-scale epidemics, *Results. Phys.*, **21** (2021), 103746. https://doi.org/10.1016/j.rinp.2020.103746

23. T. M. Chen, J. Rui, Q. Wang, A mathematical model for simulating the phase-based transmissibility of a novel coronavirus, *Infect. Disease Model.*, **5** (2020), 248—258. https://doi.org/10.1186/s40249-020-00640-3

24. P. Yarsky, Using a genetic algorithm to fit parameters of a covid-19 seir model for us states, *Math. Comput. Simulat.*, **185** (2021), 687—695. https://doi.org/10.1016/j.matcom.2021.01.022

25. Y. Fang, Y. Nie, M. Penny, Modified seir and ai prediction of the epidemics trend of covid-19 in china under public health interventions, *J. Thorac. Dis.*, **12** (2020), 165. https://doi.org/10.21037/jtd.2020.02.64
26. G. Dudek, P. Pelka, S. Smyl, A hybrid residual dilated lstm and exponential smoothing model for midterm electric load forecasting, *IEEE. T. Neur. Net. Lear.*, **33** (2021), 2879—2891. https://doi.org/10.1109/TNNLS.2020.3046629
27. S. Smyl, A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting, *Int. J. Forecast.*, **36** (2020), 75—85. https://doi.org/10.1016/j.ijforecast.2019.03.017
28. Y. Polyvianna, D. Chumachenko, T. Chumachenko, Computer aided system of time series analysis methods for forecasting the epidemics outbreaks, *EDAC* (2019). https://doi.org/10.1109/CADSM.2019.8779344
29. B. Seong, K. Lee, Intervention analysis based on exponential smoothing methods: Applications to 9/11 and covid-19 effects, *Econ. Model*, (2020). https://doi.org/10.1016/j.econmod.2020.11.014
30. H. Li, R. Zheng, Q. Zheng, W. Jiang, X. Zhang, W. Wang, et al., Predicting the number of visceral leishmaniasis cases in Kashgar, Xinjiang, China using the ARIMA-EGARCH model, *Asian. Pac. J. Trop. Med.*, **13** (2020), 81–89. https://doi.org/10.4103/1995-7645.275416
31. M. K. Lee, J. H. Paik, I. S. Na, Outbreak prediction of hepatitis a in korea based on statistical analysis and lstm network, *ICAIIC* (2020). https://doi.org/10.1109/ICAIIC48513.2020.9065082
32. S. A. Salama, M. Lavie, M. D. Buck, J. V. Damme, S. Struyf, Cytokines and serum amyloid A in the pathogenesis of hepatitis C virus infection, *Cytokine. Growth. F R*, **50** (2019), 29–42. https://doi.org/10.1016/j.cytogfr.2019.10.006
33. C. Yu, C. Xu, Y. Li, S. Yao, Y. Bai, J. Li, et al., Time Series Analysis and Forecasting of the Hand-Foot-Mouth Disease Morbidity in China Using An Advanced Exponential Smoothing State Space TBATS Model, *Infect. Drug. Resist.*, **14** (2021), 2809–2821. https://doi.org/10.2147/IDR.S304652
34. R. Ma, X. Zheng, P. Wang, H. Liu, C. Zhang, The prediction and analysis of covid-19 epidemic trend by combining lstm and markov method, *Sci. Rep.-UK*, **1** (2021). https://doi.org/10.1038/s41598-021-97037-5
35. A. Sherstinsky, Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network, *Phys. D*, **404** (2020), 132306. https://doi.org/10.1016/j.physd.2019.132306
36. Q. Wang, C. Feng, Y. Xu, H. Zhong, V. S. Sheng, A novel privacy-preserving speech recognition framework using bidirectional lstm, *Int. J. Cloud. Appl. Com.*, **9** (2020), 1–13. https://doi.org/10.1186/s13677-020-00186-7
37. Z. Karevan, J. A. Suykens, Transductive lstm for time-series prediction: An application to weather forecasting, *Neural Networks*, **125** (2020), 1—9. https://doi.org/10.1016/j.neunet.2019.12.030
38. T. Y. Kim, S. B. Cho, Predicting residential energy consumption using CNN-LSTM neural networks, *Energy*, **182** (2019), 72–81. https://doi.org/10.1016/j.energy.2019.05.230
39. X. Yan, W. Weihan, M. Chang, Research on financial assets transaction prediction model based on LSTM neural network, *Neural. Comput. Appl.*, **33** (2021), 257–270. https://doi.org/10.1007/s00521-020-04992-7

40. J. Qian, X. Qiu, X. Tan, Q. Li, J. Chen, X. Jiang, An attentive LSTM-based approach for adverse drug reactions prediction, *Appl. Intell.*, (2022), 1–15. https://doi.org/10.1007/s10489-022-03721-y

41. S. Dutta, S. K. Bandyopadhyay, Machine learning approach for confirmation of COVID-19 cases: Positive, negative, death and release, *MedRxiv*, (2020). https://doi.org/10.1101/2020.03.25.20043505

AIMS Press