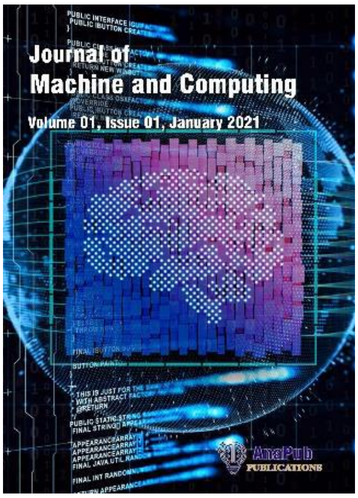
SIRD-ABiGRU-AE: A Modified Compartmental Model with Attention-Driven BiGRU Autoencoder for COVID-19 Outbreak Prediction and Hospitalizations

Parthiban M, Anna Alphy and Sreedevi B



DOI: 10.53759/7669/jmc202505055

Reference: JMC202505055

Journal: Journal of Machine and Computing.

Received 25 April 2024

Revised form 18 September 2024

Accepted 20 January 2025

Please cite this article as: Parthiban M, Anna Alphy and Sreedevi B, “SIRD-ABiGRU-AE: A Modified Compartmental Model with Attention-Driven BiGRU Autoencoder for COVID-19 Outbreak Prediction and Hospitalizations”, Journal of Machine and Computing. (2025). Doi: https:// doi.org/10.53759/7669/jmc202505055

SIRD-ABiGRU-AE: A modified compartmental model with attention-driven BiGRU Autoencoder for COVID-19 outbreak prediction and Hospitalizations

[1]Parthiban M, [1*]Anna Alphy, [2] B Sreedevi

[1]Department of Computer Science and Engineering, Faculty of Engineering and Technology,
SRM Institute of Science and Technology, Delhi–NCR Campus, Delhi - Meerut Road,
Modinagar, Ghaziabad, Uttar Pradesh, 201204, India.
[2]Department of Computer Science and Engineering
Sri Sai Ram Institute of Technology, Chennai, Tamil Nadu, 600044, India
*Corresponding Author: annaa@srmist.edu.in

Abstract: Several epidemiological studies have been undertaken using a compartmental model to predict disease spread effectively. However, knowledge about the epidemiological cycle lacks existing techniques and fails to promote the vaccines and medications that the government issues to overcome the pandemic disease. Many researchers implemented a Susceptible-Infected-Recovered-Deceased (SIRD) based compartmental approach to determine the methods emphasized by the government to eradicate the spread of COVID-19. The traditional SIRD-based compartmental model produces high prediction errors and is time-consuming. Hence, this article presents a novel Deep Learning (DL) based Attention-driven bi-directional gated recurrent unit Autoencoder (A-Bi-GRU-AE) model, which is hybridized with the SIRD model to enhance the system performance. The proposed approach is implemented in the PYTHON platform and the publicly available covid19Italy dataset is utilized for the experimental process. The proposed method obtains the overall predicted R^2 of 0.97 and time complexity of 2634.01ms.

Keywords: COVID-19, Italy, attention mechanism, bi-directional gated recurrent unit, Autoencoder, hospitalizations, compartmental models.

1. Introduction:

COVID-19 was discovered initially in Wuhan, China, in December 2019, which was then professed as a pandemic in March 2020 by the World Health Organization (WHO). Under the published real-time data by WHO, it is known that millions of people have been affected, and the mortality rates are increasing by the communicable disease [1]. COVID-19 has emerged, and some common symptoms include dry cough, appetite loss, fever and breathing difficulties, leading to complicated diseases like liver damage, septic shock and pneumonia [2, 3]. Due to this pandemic in March 2020, most countries are locked down, and strict social distancing is maintained to stop coronavirus transmission. This social distancing and lockdown aim to break the transmission chain and reduce the coronavirus. Estimating the spread over time is critical in healthcare management to protect lives and reduce the disease's social and economic consequences [4, 5].

Due to the increased contagion, the confirmed cases at the initial stage are quite increasing. As a result, a lack of ICU and respiratory equipment arose in most developing countries. The spread of COVID-19 can be eradicated with isolation beds and hospital ICUs. However, the need for isolation beds and other medical requirements is increasing in many hospitals, and the knowledge about these requirements is unknown to the governmental organization to take necessary preventive measures. To overcome this issue, an effective compartmental model is highly required to learn the daily spread of COVID-19 and other medical requirements in the hospital. Other countries like France, Belgium, New York, Japan, and South Korea report the day-to-day spread of the COVID-19 pandemic disease utilizing high effective compartmental model [6].

The compartmental model is one of the mathematical models used to calculate the count of infectious diseases by considering different compartments in an entire population [7]. During the COVID-19 pandemic, the compartmental approach predicts hospital demand and ICU utilization [8]. The common outline of compartmental modelling is that it arranges the individuals based on their disease depth and infection rate [9]. The compartmental modal considers the extra compartments for ICU and hospitalization demand [10]. The logistic functions, spreading dynamics, and standardized logistic functions are required compartmental models with infected and susceptible states [11]. The Susceptible-Infectious-Recovered (SIR) and Susceptible-Exposed-Infected-Recovered (SEIR) commonly belong to the compartmental model [12].

Mc Kendrick and Kermack introduced the compartmental model in [illegible] with SIR. Compartmental modal repeats the outbreaks of observed characteristics, like a self-limiting period. Compartmental epidemiological models depend on SEIR criteria and prolonged it for extra features consisting of ICU and health care compartments. These features are structured as 0 to 59, 60 to 79, and above 80 years of age. Several studies have been conducted using the compartmental modal, particularly in the transmission of COVID-19 in several countries, focusing on various features [13-15]. The SIR model is a type of compartmental approach consisting of three compartments: susceptible, infected and recovered [16]. In the SIR model, the epidemic spread signifies individually or transmits between susceptible-infected-recovered cases [17].

A deterministic SEIR compartmental modal is highly required to effectively calculate the spread of COVID-19. This compartmental modal mainly depends on the individuals' epidemiological status, clinical progression of COVID-19, and other intervention processes like treatment, quarantine, isolation etc. [18]. Due to the spread of COVID-19, SARS-CoV-2 creates the population's compartmental model based on the disease state level and disease awareness. The government imposed social distancing, reducing individual contact to diminish the spread of COVID-19 completely. Self-care measures are expected for each individual, including wearing masks, social distancing and hand washing. Nowadays, compartmental models are used to find epidemiological key parameters via COVID-19 clinical lessons. This compartmental modal gives the highest amount of diagnoses, time and attack rate for the highest number of COVID cases [19].

Based on the SEIR compartmental modal, the population is set to be constant with time from one compartment to another under varying infection rates. The people not present in the compartments are determined as non-infectious cases. But in the COVID-19 case, there is evidence that the people exposed in the compartment are also infectious. In this case, the people transmit the infectious diseases to the susceptible compartments. The diagnosed carriers are instantly disclosed to a hospital or isolated at home for nearly 14 days. If they are not tested, the non-diagnosed carriers with no symptoms like cough or fever can spread COVID-19 because they are not restricted in their movements or any social restrictions [20].

In recent days, several mathematical models have been proposed for understanding the dynamic progression of COVID-19. One of the best models for understanding the epidemic is a compartmental model. However, the existing models failed to provide the best approximation for the huge COVID-19 dataset. The conventional compartmental models utilize appropriate estimation approaches such as Maximum Likelihood to compute the hyper-parameters. These models usually considered time-invariant hyper-parameters and thereby reduced the prediction accuracy. Hence, the hyper-parameter should be modelled with a time-dependent characteristic to allow the model to work under varying marginal conditions. These points motivate integrating the time-dependent compartmental model with deep learning algorithms to give accurate long-term estimations. *The major contributions of this research work are listed as follows:*

- To propose a DL-integrated SIRD compartmental model by considering the time-dependent parameters to eliminate the spread of COVID-19 efficiently.
- To modify the conventional compartmental model by integrating the SIRD model with DL algorithms.
- To give an accurate long-term prediction for the Covid-19 outbreak by introducing a novel A-Bi-GRU-AE-based DL technique.
- To validate the performance of the proposed model by considering the pandemic outbreak in Italy.
- To analyze the proposed method in PYTHON and performance measures like prediction R^2 and time complexity are analyzed and compared with existing techniques.

2. Literature survey:

Some of the recent related works are listed as follows:

Keeling et al. [21] defined a different compartmental mathematical model for analyzing the spread of COVID-19 based on quarantining and age-related issues. This compartmental model forecasts the spread of an epidemic using original data based on confirmed cases. Next, difficulties with social distance were examined based on their epidemic outcomes. Lastly, discovered the key biological characteristics of COVID-19 that remain unknown under susceptibility to varying age groups and symptoms.

Ramezani et al. [22] established a new compartmental model SEAIRDQ (Susceptible-Exposed- Asymptomatic-Infectious-Recovered-Deceased-Quarantined) models for the transition of individuals between the social awareness and the suscepted compartments. The SEAIRDQ model could take the nonlinear behaviour of COVID-19 pandemic for determining the asymptomatic infections in the individuals. This model also aids in reporting the cumulative infection and death rate in various states. In addition, the SEAIRDQ model calculates an individual's current reproduction number and immunity level.

Pajaro et al. [23] determined the COVID-19 pandemic unpredictability for successfully modelling its dynamic evolution. This method aids in determining the spread of disease by training the traditional compartmental models until it returns the best prediction outcome. Here, the chemical reaction schemes were modelled using chemical master equations and solved using Monte Carlo approaches. This model was effectively used for COVID-19 prediction during pandemic conditions in medium and small-sized municipalities.

Sharov et al. [24] introduced the SIR-modified model for COVID-19 transmission to calculate the efficiency of lockdown methods during a pandemic situation. The input of this method was COVID-19 epidemiological data collected from other countries using certified information. The output parameters were considered as formation time and immunity level of that particular diseased individual. These parameters were then used as an effective indicator to determine the day-to-day analysis of the suspected cases effectively.

Rainisch et al. [25] determined the SEIR compartmental model to analyze the local pandemic transmission. This SEIR method uses the input of health care resources, case counts, and evaluates the intrusion strategies. The output includes the infection patients count, death rate, critical isolation beds, and ventilators relative to current capacity. This method shows that aggressive interventions can stop the extensive diseases and death rate from coronavirus. This SEIR method permits the fast calculation of locally applicable states and improves the outcome when the current information becomes more accurate and clear.

Wang et al. [26] defined an asymptomatic infected compartmental model by extending the proliferation rate of COVID-19. Also, this study utilized LSTM to improve prediction accuracy by updating the compartmental model parameters. Compagni et al. [27] integrated the compartmental model with a Feed-Forward Neural Network (FFNN) to increase prediction accuracy and forecast ICU occupancy.

2.1 Problem definition

In literature, some methods used feed-forward neural networks and the LSTM model to predict the patients' future trajectories. However, the feed-forward neural networks will not consider the temporal relations between the historical data. However, RNN can consider the temporal relations between the input sequences. RNN has a superior ability to encapsulate sequential information over time. Some existing approaches like LSTM have also been contemplated because RNN can't handle the gradient vanishing and long-distance dependencies problem. Hence, an effective DL model is highly required that can integrate with the compartmental model to utilize both the past and upcoming spread of the COVID-19 pandemic disease

The traditional SIRD model predicts its outcome with high error due to varying time intervals. Recently, many techniques have been integrated with the SIRD model for accurate prediction. But those techniques are highly suffered due to time complexity and error. Hence, an effective hybrid technique is required to address the drawbacks faced by the existing techniques for accurate prediction. To the best of the knowledge, the proposed method addresses all the problems arising in the existing technique and effectively provides an outstanding prediction outcome.

3. Proposed method

This paper proposes a new modified compartmental model with deep learning algorithm for predicting the COVID-19 outbreak and hospitalizations. Initially, the population is divided into four compartments: susceptible (S), infectious (I), recovered (R), and dead (d). Here, the infectious (I) compartment will consider the isolated patients at home, in the hospital and the intensive care unit. Here a dynamic transfer between each compartment is considered to show the time dependence. Then, the contact rate, recovery rate, and deceased rate will be estimated using the number of people hospitalized with symptoms, isolated patients at home and in the hospital, ICU admissions, recovered, and death data from the dataset. However, this model will not provide accurate long-term estimations. To tackle this issue, a hybridized DL-based A-Bi-GRU-AE will be integrated with **the** SIRD model for learning and correcting the estimated error created by the SIRD model. Here, the estimated results of the SIRD model will be given as input to the hybrid A-Bi-GRU-AE, and the real hospitalizations/data will be used as the target while training the A-Bi-GRU-AE model. In addition, the proposed model will introduce a similarity measure to compute the similarity between the training and the testing time series to predict the COVID-19 outbreak and hospitalization accurately**.** Also, the predicted contact and recovery rates are used to detect the epidemic progression (i.e. the reproduction number).

3.1 Prediction using hybrid SIRD with A-Bi-GRU-AE

For predicting the COVID-19 outbreak and hospitalizations, the conventional SIRD model is utilized. In the conventional SIRD model, the population M can be divided into four compartments, namely $(S),(I),(R)$ and (D) is determined based on varying time intervals v. Figure 1 determines the structure of the SIRD-based compartmental model.

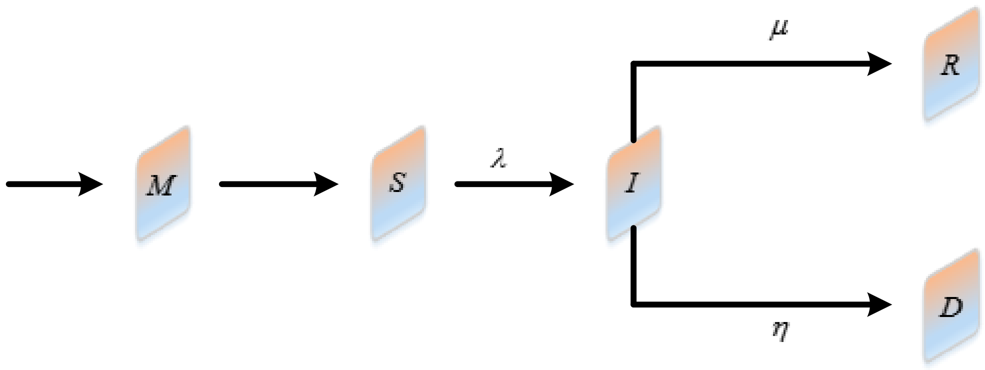


Figure 1: Structure of SIRD-based compartmental model

The mathematical interpretation of the SIRD model is explained in the upcoming section. The total population M can be formulated as,

$$M = S_v + I_v + R_v + D_v \tag{1}$$

Here, $S_v, I_v, R_v \text{ and } D_v$ indicates the suspected cases, infected cases, recovered cases and deceased cases under different time intervals, respectively. The alterations that take place in the total cases can be formulated as,

$$A_v = I_v + R_v + D_v \tag{2}$$

The obtained outcome from equation (2) always remains constant. The outcomes obtained in each compartment are depicted in detail below.

$$S_{v+1} = S - \left(\frac{\lambda SI}{M}\right) \tag{3}$$

$$I_{v+1} = I + \left(\frac{\lambda SI}{M}\right) - \mu I - \eta I \tag{4}$$

$$R_{v+1} = R + \mu I \tag{5}$$

$$D_v = D + \eta I \tag{6}$$

Here, λ determines the contact rate, μ signifies the recovery rate and η depicts the death rate. In the first phase, the susceptible (S) cases are equal to the total population M. The increasing rate of growth (ρ) for each day and the primary reproductive number (R_n) can be mathematically interpreted as,

$$\rho = \lambda - (\mu + \eta) \tag{7}$$

$$R_n = \frac{\lambda}{\mu + \eta} \tag{8}$$

When $R_n > 1(\rho > 0)$, the epidemic disease is generated and in the reverse case, the epidemic disease gets eradicated.

3.1.1 Calculation of epidemic after several measures taken by the government

Let us consider the regular change in suspected cases as $\hat{S} = S_{v+1} - S_v$. Similarly, the constant change in infected, recovered and deceased cases can be mathematically formulated as,

$$\hat{S} = \left(\frac{-\lambda SI}{M}\right) \tag{9}$$

$$\hat{I} = \left(\frac{\lambda SI}{M}\right) - \mu I - \eta I \tag{10}$$

$$\hat{R} = \mu I \tag{11}$$

$$\hat{D} = \eta I \tag{12}$$

The above interpretations are also referred to as transmission coefficients, recovery rate, and death rate.

In the past few years, the dataset $A_v = I_v + R_v + D_v$ is very less while differentiating from the total population. The total population M is very near to the S_v, then consider $\lambda = \hat{A}/I$ for a huge amount of data.

Finally, the obtained contact rate λ, recovery rate μ and, the deceased rate η can be mathematically interpreted as,

$$\lambda = \left(\frac{\hat{A}}{I}\right) \times M/(M - A) \tag{13}$$

$$\mu = \frac{\hat{R}}{I} \tag{14}$$

$$\eta = \frac{\hat{D}}{I} \tag{15}$$

Here, $\hat{A}$ represents the number of daily changes in the SIRD compartment, A indicates the total changes under different cases, $\hat{R}$ indicates the change in daily recovered cases, $\hat{D}$ denotes the changes in daily death cases.

The obtained outcome from A_v is very high, and an accurate outcome is still required to predict the compartment cases efficiently. In addition, the existing SIRD model does not show accurate number of cases due to varying time intervals. Hence, this research introduces a novel A-Bi-GRU-AE technique using the conventional SIRD model. The outcome of the conventional SIRD model is given as input to the proposed A-Bi-GRU-AE technique.

3.1.2 Proposed A-Bi-GRU-AE technique

The entire operation in the proposed model undergoes two major stages, namely offline SIRD compartmental curve library construction and online SIRD-based COVID-19 prediction estimation during the testing process. In the offline stage, embedding vectors are initially developed based on attention and skip connection (AS) with the AE model. The continuous variation of SIRD cases is evaluated by changes that have occurred in the embedding vectors during machine operations. During the training process, the curve is obtained to form the SIRD compartmental curve library.

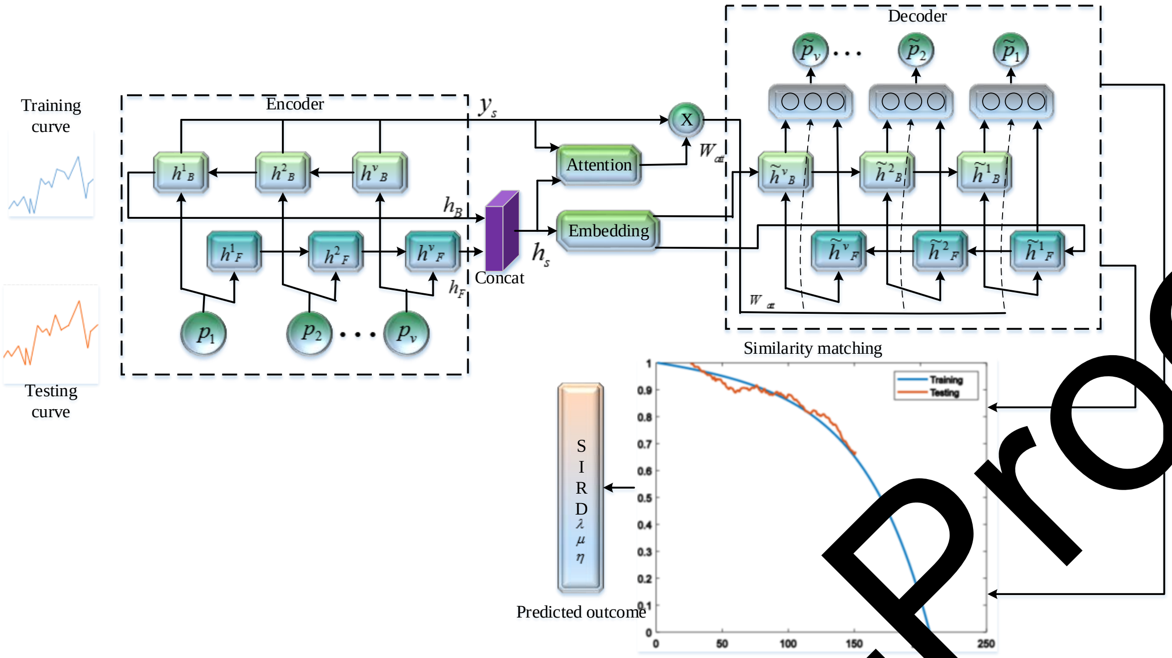


Figure 2: Architecture of A-Bi-GRU-AE model

Figure 2 depicts the architecture of the A-Bi-GRU-AE model. The proposed system works based on time series, and the curve must be smoothed to eliminate overfitting and other irreversible processes. In addition, the proposed method introduces the Linear Regression (LR) model to understand the mapping of actual and predicted outcomes from the network model. For the online phase, the test outcome is given to the trained LR model as an input for constructing the testing curve. Finally, the similarity is estimated between the offline training curve and the online testing curve for an accurate SIRD prediction.

(A). Bi-GRU with AS

The Bi-GRU consists of double GRUs having forward and backward directions that can extract time series from the input dataset. After the completion of encoding process, the hidden vector state h_s is utilized for the prediction process. The encoder having an output vector y_s gets added with h_s that can act as the input to the attention layer. Under varying time steps, the weight vector W_{att} is determined in the attention layer and it is concatenated with the output of encoder to generate an attention output vector. At the decoder phase, the attention vector output is delivered as the input to the Bi-GRU decoder. Finally, the decoded Bi-GRU outcome is obtained by the skip connection and overcomes the computational complexity by proving feature vector to the Linear Prediction (LP) layer.

Assuming the time series data as, $P = [p_1, p_2, p_3, ..., p_v]^T$ having k a number of dimensions. Also, consider that the Bi-GRU has hidden units x that encode the vector outcome. At the final stage, the hidden vector state can be mathematically formulated as,

$$\left(h_s^{1\times 2x}, y_s^{v\times 2x}\right) = G_s\left(P^{v\times k}\right)$$

$$h_s^{1\times 2x} = h_F^{1\times x} \oplus h_B^{1\times x} \qquad (16)$$

Here, $G_s\left(P^{v\times k}\right)$ depicts the summary function of the Bi-GRU encoder, $h_F^{1\times x}$ *and* $h_B^{1\times x}$ indicates the forward and backward hidden states, respectively, h_s represents the added hidden vector state from $h_F^{1\times x}$ *and* $h_B^{1\times x}$ and y_s indicates the encoder's output vector.

For evaluating the weight of the attention layer, the hidden vector state $h_s^{1\times 2x}$ and the output $y_s^{v\times 2x}$ is utilized. The initial dimension of $h_s^{1\times 2x}$ is replicated at the time v to accommodate the dimensions of $y_s^{v\times 2x}$. The weight of the attention layer $W^{1\times s}{}_{att}$ is estimated using $h_s^{1\times 2x}$ and $y_s^{v\times 2x}$ at each time step. The obtained $W^{1\times v}{}_{att}$ is then convoluted with $y_s^{v\times 2x}$ for generating $W^{1\times s}{}_{att-o}$ to extract hierarchical information during each time step. At the last stage, $W^{1\times s}{}_{att-o}$ is forwarded to the Bi-GRU's decoder unit for estimating the vector outcome and hidden state. It can be mathematically formulated as,

$$h_s^{1\times 2x} \xrightarrow{\mathrm{Re}\,plication} h_s^{v\times 2x} \tag{17}$$

$$W^{1\times v}{}_{att} = Att\left(h_s^{v\times 2x} \oplus y_s^{v\times 2x}\right) \tag{18}$$

$$W^{1\times 2x}{}_{att-o} = W^{1\times v}{}_{att} \times y_s^{v\times 2x} \tag{19}$$

$$\left(h_r^{1\times 2x}, y_r^{v\times 2x}\right) = G_r\left(W^{1\times 2x}{}_{att-o}, h^{1\times 2x}{}_s\right) \tag{20}$$

The outcome of the attention layer is then added with y_r and LP layer to decode $\tilde{p}_v$ under v a number of times for the prediction of encoder input $\tilde{P} = [\tilde{p}_1, \tilde{p}_2, \tilde{p}_3, \ldots, \tilde{p}_v]^T$. The error which is smoothened at a time v is expressed as, $E_t = \tilde{p}_t - p_t$. The AS-based Bi-GRU is then finally trained to reduce the prediction error, and it can be mathematically formulated as,

$$\Im = \frac{1}{2}\sum_{r=1}^{v}\left(\|E_t\|_1\right)^2 \tag{21}$$

Here, $\|E_t\|_1$ depicts the norm-1 operator that can intersect fast compared to a norm-2 operator. After training of AS-based Bi-GRU, the input P is compressed in the final encoded hidden state unit. If the AS-based Bi-GRU is integrated with several Bi-GRU layers, the embedding vector is generated by adding all the layers in the hidden states and can be mathematically formulated as,

$$Z_v = h^1{}_s \oplus h^2{}_s \oplus \ldots h^n{}_s \tag{22}$$

Here, $h^n{}_s$ indicates the final hidden vector state of the n^{th} layer, Z_v denotes the input time series data having embedding vector and n indicates the total Bi-GRU layers.

(B). LR-based embedding vector

This model helps to map the difference between the actual value and the predicted SIRD outcome. Assuming the failed time series as the, $P = [p_1, p_2, p_3, \ldots, p_v]^T$ and having a k number of dimensions, $p_s = [u_1, u_2 \ldots u_n]$. A stable window sliding ϖ has the data sequence

$\Psi = [\Psi_1, \Psi_2, \Psi_3, ..., \Psi_{v-\varpi+1}]$, $\Psi_x = [\Psi_x, \Psi_{x+1}, ..., \Psi_{x-\varpi+1}]^{\varpi \times n}$. The sliding window sequence is fed as the input to the AS-based Bi-GRU and hence, Ψ_x can establish an embedding vector Z_x via equation (16).

Finally, the time series having multi-dimensional features are converted into single-dimensional embedding vector series as $z = [z_\varpi, z_{\varpi+1}, z_{\varpi+2}, ..., z_v]$ that contains the details about the actual data. Assuming M as the number of cases to be predicted, thus, $z^{(b)}$ is determined as, b where, $b \in M$. The initial embedding can be mentioned as $(Z^b{}_1, Z^b{}_2, Z^b{}_3)$. $z^{(b)}$ are, must be correctly predicted, as same as values obtained in the dataset. Due to the reduction of a lifetime, the system performance gets degraded and completely alters the actual and the predicted value.

The deviation between the actual and the predicted compartmental outcomes can be mathematically formulated as,

$$D^{(b)}{}_v = \frac{1}{M} \sum_{Z \in z_{norm}} \left\| Z^{(b)}{}_v - Z \right\|_2 \tag{23}$$

Here, z_{norm} represents the normalized embedding vector, M signifies the components present in z_{norm}. The proposed system utilizes the normalization range $[0,1]$ to map the $D^{(b)}{}_v$ for generating the curve, and it can be mathematically formulated as,

$$H^{(b)}{}_v = \frac{(D^{(b)}{}_x)_{\max} - D^{(b)}{}_v}{(D^{(b)}{}_x)_{\max} - (D^{(b)}{}_x)_{\min}}, x = \varpi, \varpi + 1, ..., v^{(b)} \tag{24}$$

Here, $(D^{(b)}{}_x)_{\max}$, $(D^{(b)}{}_x)_{\min}$ depicts the maximum and minimum values having b^{th} deviation $D^{(b)}{}_v$ during a certain operational time. Based on the training of the unsupervised network, the degradation at b^{th} an instance is determined. For training the LR model, the best prediction outcome is obtained, and it can be mathematically formulated as,

$$h_v = \zeta_0 + \zeta^T p_v \tag{25}$$

Here, h_v and p_v represents the predicted value and the input reading at a time v, respectively and $\zeta = [\zeta_1, \zeta_2, ..., \zeta_n]$ represents the coefficient factors. After training the LR model, the obtained testing and training values are given as input to equation (27) to generate a particular predicted value.

(C) Similarity calculation for SIRD prediction

The proposed prediction method for the testing value is emphasized based on the similarity concept. However, the proposed technique runs longer during training and tends to reduce the prediction outcome. Let's assume that the original value is different at the training time, so the test curve is rotated with a delayed time $\aleph$ to make the training and testing curve's similarity accurately. The mathematical interpretation for the similarity calculation is depicted below:

$$Sim(b, \aleph) = \exp\left(\frac{-D(Tr_c, Te^{(b)}{}_c, \aleph)}{\vartheta} \right) \tag{26}$$

$$D\left(Tr_c, Te^{(b)}{}_c, \aleph\right) = \frac{1}{\hat{S}} \sum_{r=1}^{S} \left(\tilde{H}_s - H^{(b)}{}_{s+\aleph}\right)^2 \tag{27}$$

Here, $D\left(Tr_c, Te^{(b)}{}_c, \aleph\right)$ indicates the square of average Euclidean distance for the two curves, ϑ determines the relax factor that can measure the similarity degree under different cases, $\hat{S}$ interprets the total time taken for online process. The resting predictive value is predicted based on the b^{th} training instance and can be mathematically interpreted as,

$$\Pr ed(b, \aleph) = S_b - \hat{S} - \aleph \tag{28}$$

Here, S_b signifies the total time taken for the training process. Every testing and training instance is determined using equation (29) to generate weight similarity. The outcome having greater similarity can be mathematically interpreted as,

$$\Pr e\tilde{d} = \frac{\sum_b \sum_\aleph Sim(b, \aleph) \times \Pr ed(b, \aleph)}{\sum_b \sum_\aleph Sim(b, \aleph)} \tag{29}$$

Here, $Sim(b, \aleph) \geq \varphi \times \left(\max_{b,\aleph} Sim(b, \aleph)\right)$ which $\max_{b,\aleph} Sim(b, \aleph)$ depicts the high similarity between the training curves under varying lagging times, the parameter φ controls the training instance to be integrated with the testing instances.

4. Results and discussion

The proposed work will be implemented in the Python platform by studying the COVID-19 situation in Lombardy, Italy [28]. There are [illegible]707,743 tested cases in this dataset, which is mapped to 195,351 positive cases. For the active cases, recovered cases, hospitalizations, intensive care and death cases, 105,847, 63,120, 21,533, 2,102 and 26,384 cases are present. In addition, this dataset is separated into three categories: the total COVID cases at the regional, national, and provincial levels. A total of COVID-19 cases in Italian countries is stated from February 2020 to October 2022 under different levels respectively. Also, the performance of the proposed model will be validated by comparing the predicted results of the number of cases confirmed, death, recovered, hospitalization (ICU) and reproduction number with actual data. Also, R-squared (R^2) is a statistical measure used to measure the predicting ability of the proposed model. In addition, the integration of the DL algorithm with the SIRD model will be proved by comparing it with the conventional SIRD model. Tables 1 and 2 tabulate the proposed method's experimental details and simulation parameters.

Table 1: Experimental details of the proposed method

SYSTEM CONFIGURATION	
Device name	SST001
Full device name	SST001.seahost.local
Processor	Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz 3.40 GHz
Installed RAM	8.00 GB (7.89 GB usable)
Device ID	8591FDD2-5800-427D-BB79-151A3EB8A6AB
Product ID	00330-81495-17322-AA248
System type	64-bit operating system, x64-based processor
Pen and touch	No pen or touch input is available for this display

Table 2: Simulation parameters of the proposed method

HYPER PARAMETERS VALUES EVALUATED IN THE PROPOSED METHODOLOGY	
No. of hidden layer L	20
No. of hidden nodes h	150/128/64/50/100
Window Length W	50/25/50/100
Learning rate	0.001
Training epochs	10
Early stop	10
L2 weight	0.01
Gradient clipping	1

4.1 Performance metrics

The performance is analyzed daily to predict the COVID-19 cases under SIRD compartments. The mathematical formula for calculating the daily suspected cases is given:

$$\hat{S} = A_{v+1} - A_v \tag{30}$$

The mathematical formula for calculating the daily infected cases is given as,

$$\hat{I} = A_v - \left(R_v + D_v\right) \tag{31}$$

Also, the mathematical interpretation for calculating daily death cases is given by,

$$\hat{D} = \left(D_{v+1} - D_v\right) \tag{32}$$

For analyzing the accuracy of the proposed prediction model, the prediction coefficient R^2 is measured and can be mathematically formulated as,

$$predicted\ R^2 = \frac{\sum (Q-P)^2}{\sum (P - mean(P))^2} \tag{33}$$

Here, Q determines the confirmed and recovered cases obtained by the proposed method and P denotes the total amount of data analyzed for the prediction. If the parameter R^2 attains a negative value, then the prediction model obtains poor accuracy, and if R^2 attains a positive value, it is considered the best prediction model.

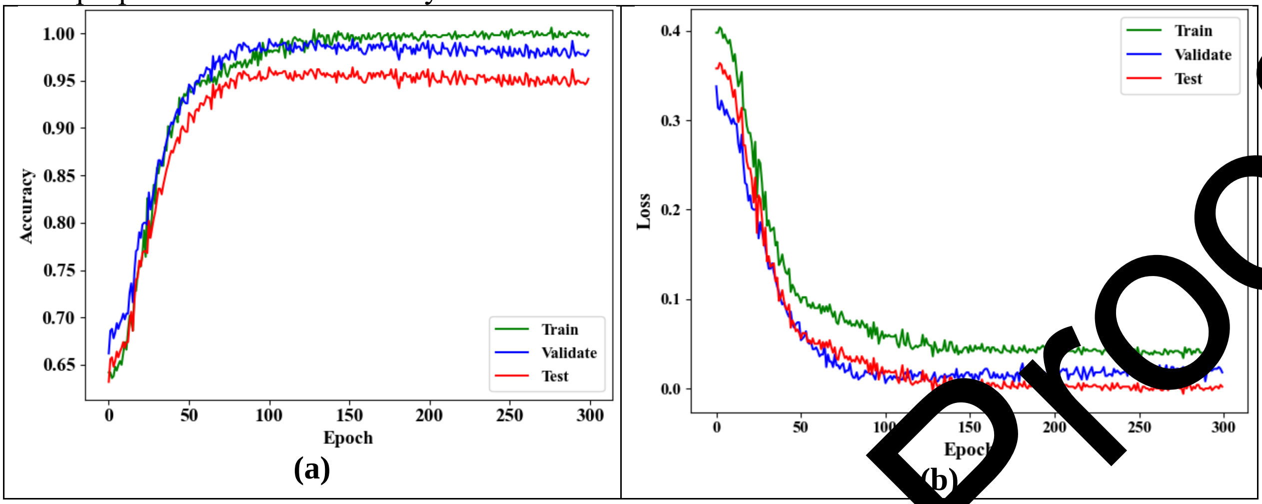
Likewise, for analyzing the error performance, MSE is measured and can be mathematically formulated as,

$$MSE = \frac{1}{l}\sum_{x=1}^{l}(P_x - \hat{P}_x)^2 \tag{34}$$

Here, l indicates the number of data used, P_x manipulates the actual value, $\hat{P}_x$ indicates the calculated value.

4.2 Comparative analysis of the proposed model with other models

This section analyses the proposed method's performance using a graphical illustration. Some other existing techniques like PDCNN, CNN, BI-GRU and GRU are also compared to prove the proposed model's efficiency.

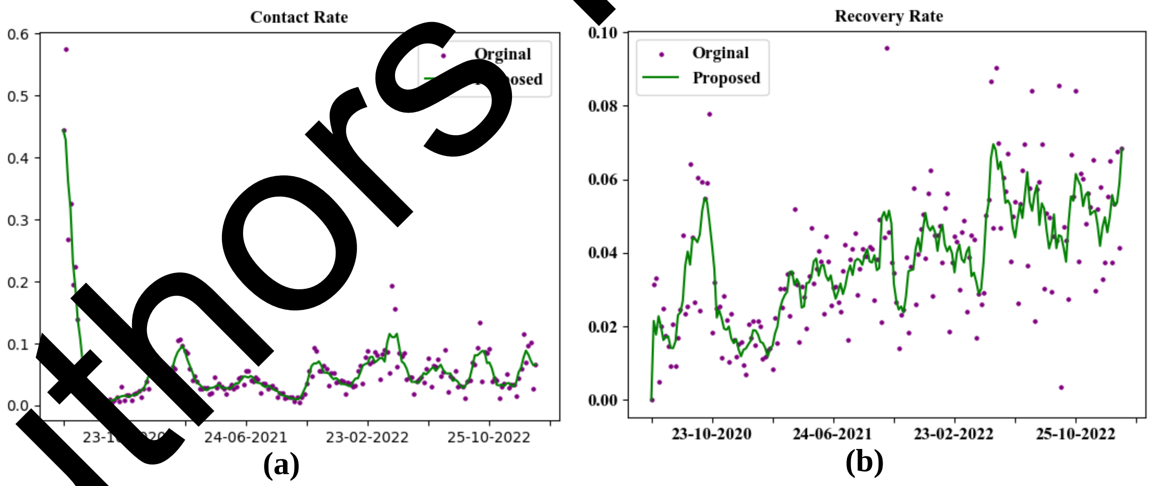


(a)

(b)

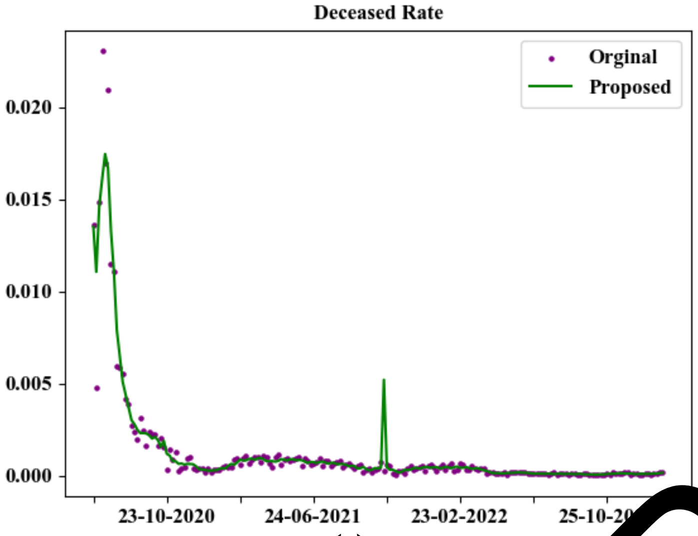
Figure 3: Accuracy and Loss curve under varying epochs, (a) Accuracy loss and (b) Loss curve

Figure 3a and 3b illustrate the accuracy and loss curve under varying epochs. The performance of the proposed model will be analyzed through training, testing and validation. From the graphical illustration, it is clear that the proposed method obtains a training accuracy of 97%, and the accuracy obtained is about 96% for the testing process. For the validation process, the accuracy obtained was about 94%. From Figure 3b, the loss obtained by the proposed method under training, testing and validation are 0.0 0.03 and 0.04, respectively.



(a)

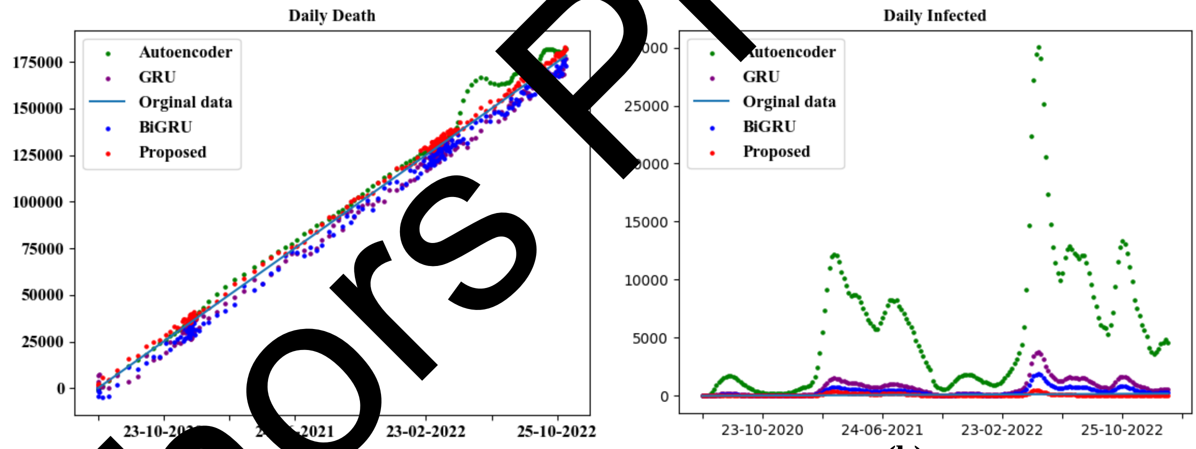
(b)



(c)

Figure 4: Analysis of (a) Contact rate, (b) Recovery rate and (c) Deceased rate

Figures 4a, 4b and 4c depict the LR plot for the contact, recovery, and deceased rates, respectively. From the graphical illustration, it is clear that the predicted value obtained by the proposed method is near to the original value. The contact rate, recovery rate and decreased rate for the proposed model are tested on a daily basis. The rate of contact, recovery and death is determined for October 2020, June 2021, February and October 2022, respectively.



(a) (b)

Figure 5: Comparative performance under (a) Daily death and (b) Daily infected cases

Figures 5a and 5b indicate the comparative performance under daily death and infected cases, respectively. The graphical plot concludes that the proposed method obtains almost similar outcome compared to the original value. In contrast, the existing techniques continuously show a random outcome rather than the actual value. The daily death and infected cases are emphasized for October 2020, June 2021, February and October 2022, respectively. For 24th June 2021, the total original death cases are considered 127362, and the proposed predictive model effectively predicted 127361 death cases. For the infected cases, on 24th June 2021, the original cases are given as 68619, and the proposed hybrid model correctly predicts the total 68618 infected cases.

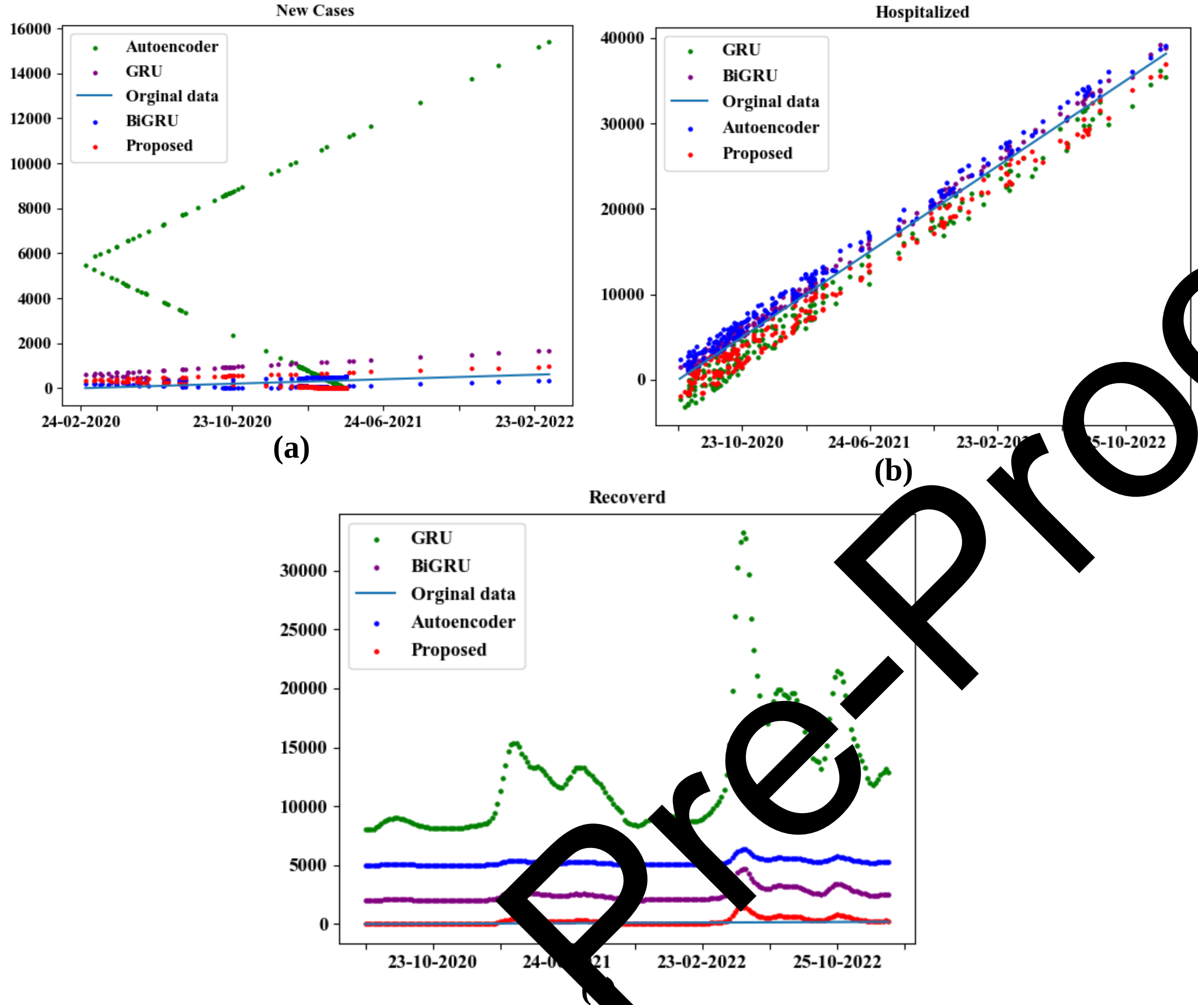


Figure 6: Comparative performance under (a) New cases, (b) Hospitalized and (c) Recovered cases

Figures 6a, 6b, and 6c illustrate the comparative performance under new, hospitalized, and recovered cases, respectively. The graphical interpretation gives a close result compared to other conventional models for the proposed method. The new cases, hospitalized and recovered cases are determined for October 2020, June 2021, February and October 2022, respectively. By 23rd October 2020, the total original new cases are determined as 186,002, and the proposed model correctly predicts a total of 186,001. For hospitalization, a total of 38507 people were hospitalized on 23rd October 2020, and a total of 38503 were corrected predicted. For recovered cases, a total of 58449 people are given in the dataset, and 58451 recovered cases are correctly predicted. The proposed technique is also compared with existing techniques, and the predicted outcome of these techniques is completely away from the actual value. However, the SIRD compartment model is a highly time-dependent process and needs to train for a longer time to get an accurate prediction. The existing techniques cannot be supported for multiple varying periods, and they suffer greatly from the gradient vanishing problem.

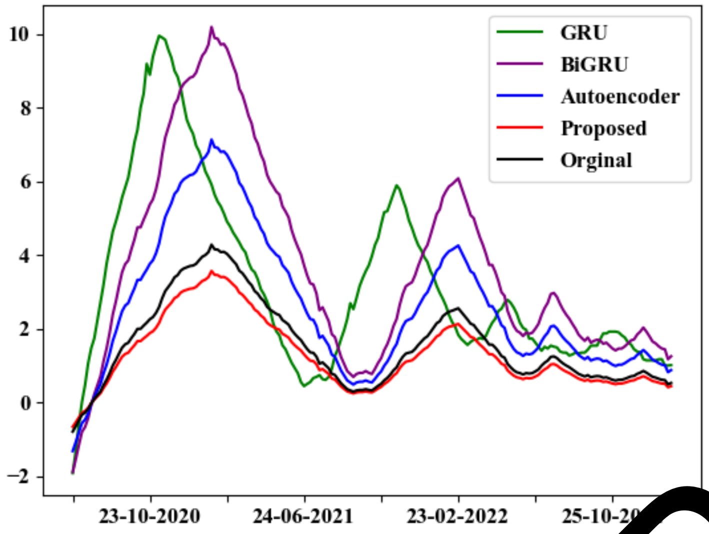


Figure 7: Performance comparison for analyzing reproduction number

Figure 7 signifies the performance comparison for analyzing the number of reproductions. The proposed method obtains near to the original value from the graphical illustration. The reproduction number is determined for varying months and years as October 2020, June 2021, February, and October 2022, respectively. The increase in the growth of reproduction number starts to reduce from February 2022 and maintains the constant outcome throughout the year 2022. During the beginning of the COVID-19 outbreak in 2020, the number of reproductions grew and rapidly fell to the 0th position in 2021. The proposed hybridized predictive model is also compared with multiple traditional techniques and proves the efficiency of the hybrid model.

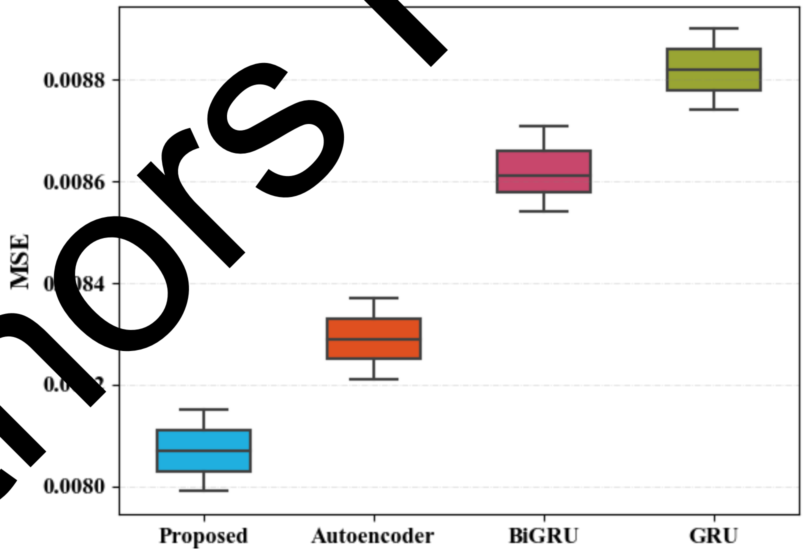


Figure 8: Error performance under different techniques

Figure 8 contemplates the error performance under different techniques. The proposed model obtains a low error from the graphical manipulation compared to other conventional techniques. The conventional AE, Bi-GRU, GRU and the proposed hybrid model obtain the MSE of 0.0082, 0.0089, 0.00861, 0.0088 and 0.008, respectively. From the experimental outcome, the existing techniques show major difference between the predicted and the original value. Table 3 tabulates the outcome of $predicted\ R^2$ and time complexity. Here,

$predicted\ R^2$ is one of the effective performance metrics for analyzing the effectiveness of the proposed predictive model. In addition, the time complexity of the proposed method is also analyzed and compared with different traditional techniques.

Table 3: Outcome of $predicted\ R^2$ and time complexity

Performance measures	Proposed	AE	Bi-GRU	GRU
$predicted\ R^2$	**0.97**	0.89	0.78	0.77
Time complexity (ms)	**2634.01**	6069.17	7483.31	10212.89

5. Conclusion

For accurate prediction of the spread of COVID-19 outbreaks and hospitalizations, the traditional SIRD compartment model is not applicable for training with huge data. The conventional SIRD model splits the compartments into four parts, and stable transmission is determined based on varying time intervals. Using suspected, infected, recovered, and death cases, the rate of contact, recovered, and deceased are predicted. However, the traditional compartment model is considered a time-consuming process and cannot handle daily changes in COVID cases. This research brought a novel hybridized A-Bi-GRU-AE-based DL algorithm that aids the Italian government in taking necessary interventions and future decisions to deal with the pandemic. The proposed method obtains the overall predicted R^2 of 0.97 and time complexity of 2634.01ms. The proposed method's main advantage is that it can process huge datasets with low time complexity. Despite this, the proposed method utilizes a single dataset for the whole process. In future, the researchers need to focus on utilizing the proposed method for processing multiple datasets to eradicate the spread of COVID-19 effectively.

Reference

1) Abrams, Steven, James Wambua, Eva Santermans, Lander Willem, Elise Kuylen, Pietro Coletti, Pieter Libin et al. "Modelling the early phase of the Belgian COVID-19 epidemic using a stochastic compartmental model and studying its implied future trajectories." *Epidemics* 35 (2021): 100449.
2) Yang, Li, Shasha Liu, Jinyan Liu, Zhixin Zhang, Xiaochun Wan, Bo Huang, Youhai Chen, and Yi Zhang. "COVID-19: immunopathogenesis and Immunotherapeutics." *Signal transduction and targeted therapy* 5, no. 1 (2020): 1-8.
3) Kandel, Nirmal, Stella Chungong, Abbas Omaar, and Jun Xing. "Health security capacities in the context of COVID-19 outbreak: an analysis of International Health Regulations annual report data from 182 countries." *The Lancet* 395, no. 10229 (2020): 1047-1053.
4) Sameni, Reza. "Mathematical modeling of epidemic diseases; a case study of the COVID-19 coronavirus." *arXiv preprint arXiv:2003.11371* (2020).
5) Di Domenico, Laura, Giulia Pullano, Chiara E. Sabbatini, Pierre-Yves Boëlle, and Vittoria Colizza. "Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies." *BMC medicine* 18, no. 1 (2020): 1-13.
6) Jen, Grace Hsiao-Hsuan, Shey-Ying Chen, Wei-Jung Chang, Chiung-Nien Chen, Amy Ming-Fang Yen, and Ray-E. Chang. "Evaluating medical capacity for hospitalization and intensive care unit of COVID-19: A queue model approach." *Journal of the Formosan Medical Association* 120 (2021): S86-S94.

7) Bousquet, Arthur, William H. Conrad, Said Omer Sadat, Nelli Vardanyan, and Youngjoon Hong. “Deep learning forecasting using time-varying parameters of the SIRD model for Covid-19.” *Scientific Reports* 12, no. 1 (2022): 1-13.
8) Bollon, Jordy, Matteo Paganini, Consuelo Rubina Nava, Nello De Vita, Rosanna Vaschetto, Luca Ragazzoni, Francesco Della Corte, and Francesco Barone-Adesi. “Predicted effects of stopping COVID-19 lockdown on Italian Hospital demand.” *Disaster medicine and public health preparedness* 14, no. 5 (2020): 638-642.
9) Thompson, Robin N. “Epidemiological models are important tools for guiding COVID-19 interventions.” *BMC medicine* 18, no. 1 (2020): 1-4.
10) Chowdhury, Rajiv, Kevin Heng, Md Shajedur Rahman Shawon, Gabriel Goh, Daisy Okonofua, Carolina Ochoa-Rosales, Valentina Gonzalez-Jaramillo et al. “Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries.” *European journal of epidemiology* 35, no. [illegible] (2020): 389-399.
11) Liu, Xianbo, Xie Zheng, and Balakumar Balachandran. “COVID-19: data-driven dynamics, statistical and distributed delay models, and observations.” *Nonlinear Dynamics* 101, no. 3 (2020): 1527-1543.
12) Singh, Sudhansu Sekhar, and D. K. Mohapatra. “Predictive analysis for covid-19 spread in india by adaptive compartmental model.” *medRxiv* (2020).
13) Eubank, Stephan, I. Eckstrand, B. Lewis, Srinivasan Venkatramanan, Madhav Marathe, and C. L. Barrett. “Commentary on Ferguson, et al., “Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand”.” *Bulletin of mathematical biology* 82, no. 4 (2020): 1-7.
14) Sjödin, Henrik, Anders F. Johansson, Åke Brännström, Zia Farooq, Hedi Katre Kriit, Annelies Wilder-Smith, Christofer Åström, Johan Thunberg, Mårten Söderquist, and Joacim Rocklöv. “COVID-19 healthcare demand and mortality in Sweden in response to non-pharmaceutical mitigation and suppression scenarios.” *International journal of epidemiology* 49, no. 5 (2020): 1443-1453.
15) Gatto, Andrea, Gabriele Accarino, Valeria Aloisi, Francesco Immorlano, Francesco Donato, and Giovanni Aloisio. "Limits of Compartmental Models and New Opportunities for Machine Learning: A Case Study to Forecast the Second Wave of COVID-19 Hospitalizations in Lombardy, Italy." In *Informatics*, vol. 8, no. 3, p. 57. MDPI, 2021.
16) Tolles, Juliana, and ThaiBinh Luong. “Modeling epidemics with compartmental models.” *Jama* 323, no. 24 (2020): 2515-2516.
17) Weissman, Gary E., Andrew Crane-Droesch, Corey Chivers, ThaiBinh Luong, Asaf Hanish, Michael Z. Levy, Jason Lubken et al. “Locally informed simulation to predict hospital capacity needs during the COVID-19 pandemic.” *Annals of internal medicine* 173, no. [illegible] (2020): 21-28.
18) Rafiq, Danish, Asiya Batool, and M. A. Bazaz. “Three months of COVID-19: A systematic review and meta-analysis.” *Reviews in medical virology* 30, no. 4 (2020): e2113.
19) Teslya, Alexandra, Thi Mui Pham, Noortje G. Godijk, Mirjam E. Kretzschmar, Martin CJ Bootsma, and Ganna Rozhnova. “Impact of self-imposed prevention measures and short-term government-imposed social distancing on mitigating and delaying a COVID-19 epidemic: A modelling study.” *PLoS medicine* 17, no. 7 (2020): e1003166.
20) Leontitsis, Alexandros, Abiola Senok, Alawi Alsheikh-Ali, Younus Al Nasser, Tom Loney, and Aamena Alshamsi. “Seahir: A specialized compartmental model for covid-19.” *International journal of environmental research and public health* 18, no. 5 (2021): 2667.

21) Puttaswamy, B. S., and N. Thillaiarasu. "Fine DenseNet based human personality recognition using english hand writing of non-native speakers." Biomedical Signal Processing and Control 99 (2025): 106910.

22) Keeling, Matt J., Edward M. Hill, Erin E. Gorsich, Bridget Penman, Glen Guyver-Fletcher, Alex Holmes, Trystan Leng et al. “Predictions of COVID-19 dynamics in the UK: Short-term forecasting and analysis of potential exit strategies.” *PLoS computational biology* 17, no. 1 (2021): e1008619.

23) Ramezani, Somayeh Bakhtiari, Amin Amirlatifi, and Shahram Rahimi. “A novel compartmental model to capture the nonlinear trend of COVID-19.” *Computers in Biology and Medicine* 134 (2021): 104421.

24) Pájaro, Manuel, Noelia M. Fajar, Antonio A. Alonso, and Irene Otero-Muras. “Stochastic SIR model predicts the evolution of COVID-19 epidemics from public health and wastewater data in small and medium-sized municipalities: A one year study.” *Chaos, Solitons & Fractals* (2022): 112671.

25) Ashwin Shenoy, M., and N. Thillaiarasu. "Enhancing temple surveillance through human activity recognition: A novel dataset and YOLOv4-ConvLSTM approach." Journal of Intelligent & Fuzzy Systems Preprint (2023): 1-16.

26) Sharov, Konstantin S. “Creating and applying SIR modified compartmental model for calculation of COVID-19 lockdown efficiency.” *Chaos, Solitons & Fractals* 141 (2020): 110295.

27) Rainisch, Gabriel, Eduardo A. Undurraga, and Gerardo Chowell. “A dynamic modeling tool for estimating healthcare demand from the COVID19 epidemic and evaluating population-wide interventions.” *International Journal of Infectious Diseases* 96 (2020): 376-383.

28) Ravi Prasad, M., and N. Thillaiarasu. "Multichannel EfficientNet B7 with attention mechanism using multimodal biometric-based authentication for ATM transaction." Multiagent and Grid Systems 20.2 (2024): 89-108.

29) Wang, Liya, Yaxun Dai, Renzhuo Wang, Yuwen Sun, Chunying Zhang, Zhiwei Yang, and Yuqing Sun. “SEIARN: Intelligent Early Warning Model of Epidemic Spread Based on LSTM Trajectory Prediction.” *Mathematics* 10, no. 17 (2022): 3046.

30) Delli Compagni, Riccardo, Zhao Cheng, Stefania Russo, and Thomas P. Van Boeckel. “A hybrid Neural Network-SEIR model for forecasting intensive care occupancy in Switzerland during COVID-19 epidemics.” *PloS one* 17, no. 3 (2022): e0263789.

31) https://github.com/RamiKrispin/covid19Italy