Dynamic SEIR-GNN Using Individualized Data For Risk Prediction During The Pandemic

Jemina J[1,*], Sheeba Singh[2]

[1] Department of Mathematics, Noorul Islam Centre for Higher Education, Kumarakovil, Tamil Nadu, India
[2] Department of Mathematics, Noorul Islam Centre for Higher Education, Kumarakovil, Tamil Nadu, India
* Correspondence: Email:cyberphd85@gmail.com
journals856@gmail.com
* ORCID : 0009-0009-0543-5920

Abstract

Predicting the impact of a pandemic on a population is a question that has been debated for decades and remains unresolved. In our proposed study, we adopted a strategy that leverages the mechanistic understanding of compartmental models and the high adaptability of deep learning models to design a more effective mathematical model for pandemic prediction. This mathematical model is designed to predict the severity of future pandemic outbreaks based on geographic location and the type of hotspots individuals access daily, incorporating both human transmission and environmental transmission variables to enhance prediction accuracy. The proposed model uses a SEIR compartmental model with a multilayered GNN algorithm to predict the model parameters. The proposed multilayered GNN comprises individual agents, clusters based on the type of importance, and significant geographical clusters, each representing individual nodes and layers within the GNN. The edge weight between nodes is updated based on risk scores, which are calculated based on the likelihood of an agent carrying a pathogen while interacting with that region. Using this model, it is possible to predict the outcome of a pandemic by changing various parameters and conditions. This adaptability will be invaluable when planning for long-term pandemic predictions, particularly when the initial parameters of transmission rates and recovery rates for individuals remain unclear. The proposed model yields better prediction accuracy, with a Root Mean Square Error (RMSE) value of 2.85 and a Mean Absolute Error (MAE) of 2.49, outperforming standalone compartmental models. Policymakers and governmental agencies can use this hybrid model to frame various control measures during the pandemic.

Keywords: Compartmental models, SEIR, pandemic prediction, hybrid models, mechanistic models.

INTRODUCTION

Daniel Bernoulli introduced the concept of predicting an epidemic using a mathematical model in 1766. He introduced a life table-based model to find the mortality during the smallpox pandemic in England [1]. After that, the epidemiological model has evolved from Stochastic Models to present-day SEIR models [2]. These models are developed under the assumption of homogeneity, which means that every population group interacts with others in a structured manner with constant transmission rates. They lack Stochasticity in their form and can be highly static, reducing the real-life nuances [3][4]. The current mathematical modelling for pandemic prediction does not account for heterogeneity in population modeling. These models disregard social and individual behavior patterns entirely to predict diseases accurately. They require more individualized and detailed data collection to make more accurate predictions [5]. In some applications, SEIR models are further extended to provide detailed predictions, including both "Hospitalized" and "Deceased" data. Epidemiological models are optimized by incorporating machine learning models, which can forecast the risk of infection and potential outcomes, thereby helping to understand intervention strategies [6] better. These ML models excel at extracting auxiliary information from disparate data sources, which is beneficial for real-time prediction of potential outbreaks [7]. However, existing ML and deterministic differential equation-based models often lack the inclusion of socio-economic and demographic factors, which play a massive role in determining the outbreak [8].

The pandemic prediction models are primarily divided into two groups: stochastic models, which account for randomness in pandemic growth to predict the future, and deterministic models, which assume a predictable pattern in pandemic growth. At the same time, deterministic models have a set of predefined functions and rules that provide a clear understanding of the pandemic's behavior. Deterministic models can be explained mathematically. They are suitable for larger population groups, but they often lack randomness and miss some nuances that apply to each new pandemic [9]. This type of model does not consider individual dynamics before predicting the pandemic movements. In those deterministic models, we can categorize them into statistical models, which use previous data to predict future outcomes. However, this type of statistical pandemic prediction does not aid in detecting novel pandemics, as it lacks sufficient evidence to do so. Compartmental models are widely used mathematical models that can be employed to derive nuanced information based on simple differential equations. These types of models are suitable for predicting longer time durations and are theoretically relevant. One of the primary examples of the compartmental model is the SEIR model, which categorizes the entire population into four distinct groups: susceptible, exposed, infected, and recovered. We can further categorize the system to include vaccinated and immune individuals, depending on the type of pandemic. The newly popularised pandemic prediction method uses machine learning algorithms to predict pandemic progression [10].

All machine learning models require a large amount of labeled data and often overlook contact patterns between individuals. Many simple tree-based algorithms lack sensitivity to the complex dynamics of a pandemic. Even algorithms with complex architectures, such as convolutional and recurrent networks, face the problem of hyperparameter optimization. The most common type of machine learning model used for epidemiological prediction is Graph Neural Networks (GNN). The GNN algorithm uses nodes and edges in a graph to represent the connection pattern between individuals. This algorithm will help better understand the transmission pattern of highly infectious diseases [11]. These algorithms often lack significant insights into disease dynamics, which can hinder the prediction results. There is also a lack of historical data for novel pandemics, which can be a problem while using GNN as a standalone algorithm for pandemic prediction. Therefore, using machine learning algorithms in conjunction with compartmental models can provide a better foundation and improve adaptability and accuracy [12].

Implementing machine learning algorithms, such as random forest, is more complicated than networks like LSTM, which can increase the model's capability to predict complex dependencies in the pandemic. Instead of relying on historical pandemic data, machine learning algorithms can utilize various other essential factors, such as climate, seasonality, population demographics, and the availability of healthcare services, to predict multiple compartmental model results. Using this type of neural network-based prediction model, we can make accurate predictions at a more hyperlocal level, ranging from national to local progress of a pandemic. Therefore, we are incorporating a hybrid approach that utilizes a multilayered graph neural network (GNN) algorithm for predicting variables in SEIR models. The study also introduced a novel SEIR model that considers both environmental transmission and mobility factors before estimating the results. The entire population in the region is divided into a more hyperlocal arrangement based on activities such as transit, grocery stores for purchasing basic goods, and employment in essential jobs. The entire population is segmented, and a neural network-based risk assessment model is employed to estimate the risk of movement for each individual based on their demographics, as illustrated in Figure 1 below.

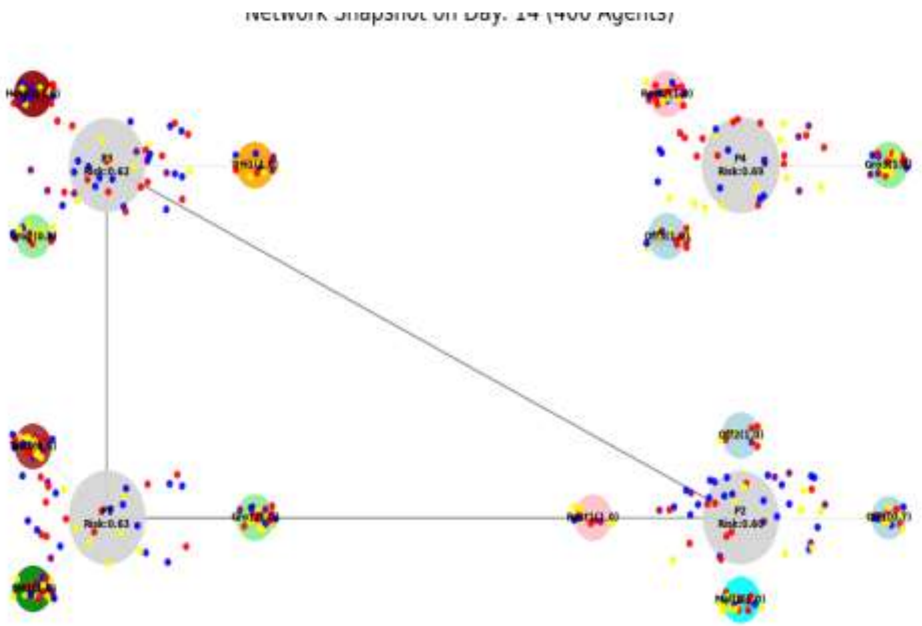

Figure 1: An illustration of risk prediction using a multilayered GNN algorithm.

Literature review

Many existing SEIR models often consider people in a simple homogeneous group without any random patterns. This reduces the accuracy of pandemic predictions as people's behavior changes under various circumstances. The study by Hou, Y., & Bidkhori, H. (2024) utilized factors such as health conditions and social activity levels for each individual to accurately predict the susceptibility level. This, along with the prediction, provides better insights into prioritizing the correct vaccines for individuals [13]. Instead of relying on social patterns of demographics, hereditary traits can be used for more accurate prediction. The study conducted by Soulaimani et al. (2024) utilized fractional derivatives with memory functions to analyze the historical values of the model's predictions. The model performs well for pandemics with a singular mode of transmission but lacks accuracy in predicting multidimensional variations in pandemic detections [14].

Instead of historical variables, authors Ushirobira, R., & Efimov, D. (2020) used the impact of social distancing and societal conditions on SEIR models. This uses novel factors to increase the overall prediction accuracy compared to existing SEIR models [15]. To incorporate uncertainty into the model, Khastan et al. (2020) used fuzzy parameters in the model. This reduces the vagueness, but still has problems with real-world uncertainty. This method includes granular derivatives into the system, which can increase the overall computation time taken for these methods [16]. The author Jin Wang (2020) analyzed multiple compartmental models and fractional models, concluding that simplifying complex pandemic patterns into these simple equations can reduce the ability to predict pandemics for a more extended period. He also argued that using only compartmental models can make them dependent on the accuracy of parameters and input data, which are not constant for many real-world pandemic movements. Therefore, using a hybrid method of machine learning models with compartmental models can be a better idea [17].

The challenges faced by compartmental models are addressed by researchers Ahmadini, A. A. H. et al. (2025). In this paper, a Recurrent Neural Network (RNN) is implemented in the SEIR model to predict the complex parameters. The proposed algorithm utilizes real-world data to train the model and generate output, resulting in improved prediction accuracy compared to other methods. The primary limitation of the paper is the reliance on pre-trained RNN models, which restricts the performance of the proposed method [18]. Cheng addresses the noise dataset problem of all existing machine learning-based modeling, C. et al. (2025). In this research, a hybrid model was developed by combining the SEIRV model and the

Dynamic Informed Neural Network (DINN) to predict the progression of pandemics. They utilized the loss function derived from SEIR-generated data to predict the model's performance. The proposed model yields better RMSE and MSA values compared to using compartmental models alone [19]. The temporal dynamics of these data are well addressed by Zheng, Y. et al. (2020); in this research, the RNN algorithm is used in combination with the SEIR model to identify patterns in the data. This study emphasized the use of mobility and geographic proximity values to determine the probability of infection. These types of hybrid models can outperform SEIR or machine learning models for short-term predictions, but they often struggle to predict data for long-term scenarios [20]. The above studies found that using hybrid methods, which combine compartmental models and machine learning algorithms, can yield better predictive accuracy than standalone models or machine learning algorithms.

Methodology

We employ fractional order analysis on the SEIR model, utilizing mobility and demographic values to determine the SEIR values for a given population. The population dynamics and other vital factors are calculated using the multilayered GNN algorithm. The model is trained and tested using real-world databases. The GNN algorithm will predict the parameters for the SEIR model. The SEIR model is validated using the Ulam–Hyers theorem, Banach contraction principle, and Krasnoselskii fixed-point theorem.

Give a ranking or weightage method for each factor for a pandemic. The weightage should be based on its impact on the pandemic outcome.

The proposed model considers both the human mode of transfer and transfer through environmental pathways, utilizing contact types, mobility, and exposure to environments as modulating factors. We introduced the total load of environmental pathogens values. E_{env}, which can be transmitted through the air, water, or by the bite of mosquitoes. The total value of susceptibility and exposure can be adapted to the values of direct exposure (βSI) and exposure through contact with the environment. $\beta_{env} SE_{env}$ Exposure values. These values are then substituted into the existing SEIR model equations to obtain the formulas below ((1) to (4)).

$$\frac{dS}{dt} = N - r_t M_h SI - r_{env} M_w SE_{env} + \delta R \tag{1}$$

$$\frac{dE}{dt} = r_t M_h SI + r_{env} M_w SE_{env} - \sigma E \tag{2}$$

$$\frac{dI}{dt} = \sigma E - (\gamma_r + \nu) I \tag{3}$$

$$\frac{dR}{dt} = \gamma_r I - \delta R \tag{4}$$

M_h: indicate the movement from the hot zone.
Λ: represents the total susceptible population.
σ\sigmaσ: Rate at which exposed individuals become infectious (inverse of incubation period).
γ_rRecovery rate from infection.
δ: Rate of immunity loss.
λI\lambda_IλI: Rate at which infected individuals shed pathogen into the environment
δE\delta_EδE: Death rate of the pathogen in the environment (from the metapopulation model).
λS\lambda_SλS: Rate at which susceptible individuals take up the pathogen from the environment

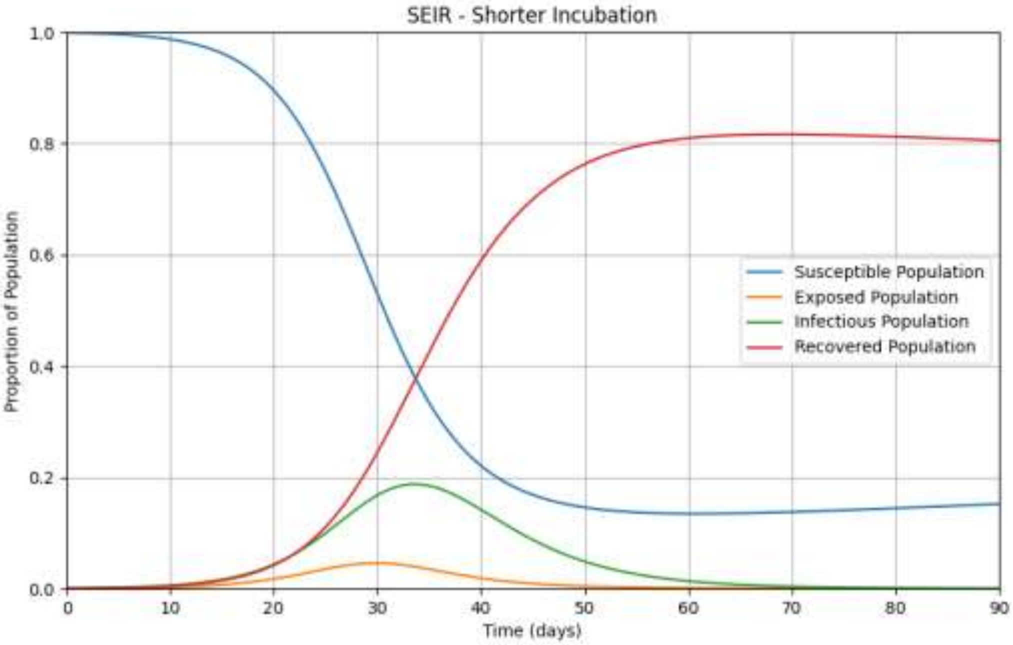

Figure 2: Model equation output for the proposed model with median parameter values.

Stability analysis of the proposed SEIR model:

The stability analysis will help identify which parameter to focus on and what control measures can be more effective in disease prevention. This metric will predict when a pandemic will die out and how long it can persist in a population.

In disease-free equilibrium, the value of infection I=0 is made zero, and at time 0, the value of every equation becomes $E^0 = 0, I^0 = 0$, $R^0 = 0$ The susceptibility value will be $S^0 = N$ Considering the population at the start is constant and does not include the birth rate. The model predicts that even when the infection rate in humans is zero, if $E_{env} > 0$The possibility of infection and exposure still exists and can affect the population. In this case, the disease-free equilibrium does not hold, as the value of exposure becomes as shown in Equation (5) below.

$$E = r_{env} M_w S E_{env} \tag{5}$$

Making the disease-free equilibrium unstable.

To estimate the rate of infections that an individual can cause, we evaluate the reproduction number using the Next-Generation Matrix Method. Using the proposed SEIR model formula, the fundamental transmission matrix can be represented as follows (6).

$$X = \begin{bmatrix} S(t) \\ E(t) \\ I(t) \\ R(t) \end{bmatrix} \tag{6}$$

Using the disease-free equilibrium estimate as the ground, evaluate the Jacobian matrix values for susceptible and exposed population groups. The reproduction number for the proposed model can be calculated based on the following equations.

$$F = \begin{bmatrix} 0 & r_t M_h N \\ 0 & 0 \end{bmatrix} \tag{7}$$

$$\frac{dE}{dt} = -\sigma E \tag{8}$$

$$\frac{dI}{dt} = \sigma E - (\gamma_r + \nu) I \tag{9}$$

$$U = \begin{bmatrix} \sigma & 0 \\ -\sigma & \gamma_r + \nu \end{bmatrix} \tag{10}$$

$$U^{-1} = \frac{1}{\sigma(\gamma_r+\nu)} \begin{bmatrix} \gamma_r + \nu & 0 \\ \sigma & \sigma \end{bmatrix} \tag{11}$$

$$FU^{-1} = \begin{bmatrix} 0 & r_t M_h N \\ 0 & 0 \end{bmatrix} \cdot \frac{1}{\sigma(\gamma_r+\nu)} \begin{bmatrix} \gamma_r + \nu & 0 \\ \sigma & \sigma \end{bmatrix} \tag{12}$$

$$FU^{-1} = \frac{1}{\sigma(\gamma_r+\nu)} \begin{bmatrix} \sigma r_t M_h N & \sigma r_t M_h N \\ 0 & 0 \end{bmatrix} \tag{13}$$

Using the eigenvalues of the matrix, we can calculate the introductory reproduction rate of the pandemic. R_o.

$$R_0 = \frac{r_t M_h N}{(\gamma_r+\nu)} \tag{14}$$

The spreading of diseases can be impacted by the R_0If the value is higher than 1, then the chance of diseases spreading is more rapid. The Jacobian matrix for the proposed model can be characterized by the equation below (15).

$$J_0 = \begin{bmatrix} 0 & 0 & -r_t M_h N & \delta \\ 0 & -\sigma & r_t M_h N & 0 \\ 0 & \sigma & -(\gamma_r + \nu) & 0 \\ 0 & 0 & \gamma_r & -\delta \end{bmatrix} \tag{15}$$

To estimate the asymptotic stability, we analyze whether there is a presence of negative real numbers for all values of the reproduction numbers that are less than 1. The eigenvalues are calculated and analyzed with equation (16).

$$\lambda_1 = \sigma + \gamma_r + \nu$$
$$\lambda_2 = \sigma(\gamma_r + \nu)(1 - R_0) \tag{16}$$

For all the values of $R_0 < 0$, the value of $\lambda_1$and λ_2 seems to have negative absolute values as solutions, proving the asymptotic stability of the proposed model.

Multilayered GNN

The algorithm comprises multiple levels of graph structure, including individuals, hyperlocal clusters, and geographic areas, all represented as nodes. The nodes are assigned in a way that individuals or agents are assigned as nodes at the lower end of the layer, and they are grouped into a cluster of hyperlocal locations. These locations are decided based on mobility and agents' behavior in those areas. The connection between agents and these hyper-local clusters is marked as edges. The hyper-local nodes are heterogeneous and are located in a specific geographic region.

To understand mobility, a simple interaction model using Gaussian Mixture Models (GMMs) is employed. The probability density function can be represented as a simple equation (17).

$$p_i = \sum_{k=1}^{N} M_k N \tag{17}$$

WhereP_i represents node 'i' probability density function, $M_k$as the mixing coefficient, and the normal distribution is represented as N.

The weight of edges in the clusters is updated using risk score values, which are calculated based on demographics, health data, frequency of movement, and the geographic cluster in which they are currently located. The weight of the edges will be updated using temporal decay values. The risk value of an individual moving from one cluster to another can be calculated using the formula (18).

$$R_{c1-c2} = \alpha(\rho_{c2}.\theta_{c2}) + \beta(\delta_s(1 - P_i) + \delta_i T_i) + \tau f_{c1} w_{c1-c2} \tag{18}$$

Where $\rho_{c2}$is the density of cluster c2, $\theta_{c2}$is the weight of location, δ_s, δ_i susceptibility levels, $P_i$is the stringency level, $and\ T_i$ is the transmission potential of a person. In the interaction-based formula, f_{c1} represents interaction frequency in C1 and edge weight between C1 and C2 w_{c1-c2}. The assigned weight values are represented as α, β, τ.

The individual, cluster, and regional node values are then fed to the message passing layer of the GNN algorithm, as shown in Figure 3 below. Each layer is designed for a specific purpose. The bottom layer is used to understand local spatial patterns, and the top layers are used to create a global graph representation.

The input variables are fed to $I_{i,c,}$ and inputs, and the node representations are understood. Then, in the next layer, the node values represented from all regional layers were aggregated to estimate the temporal representation of the data further. The SEIR variables are classified in the classification layer of the GNN and can be represented in the figure below. The first layer processes the input feature vectors, and the output of the first layer acts as the input, along with embedding outputs from all other geographic locations. This forms a multilayered structure analysis, calculating the parameters using historical values.

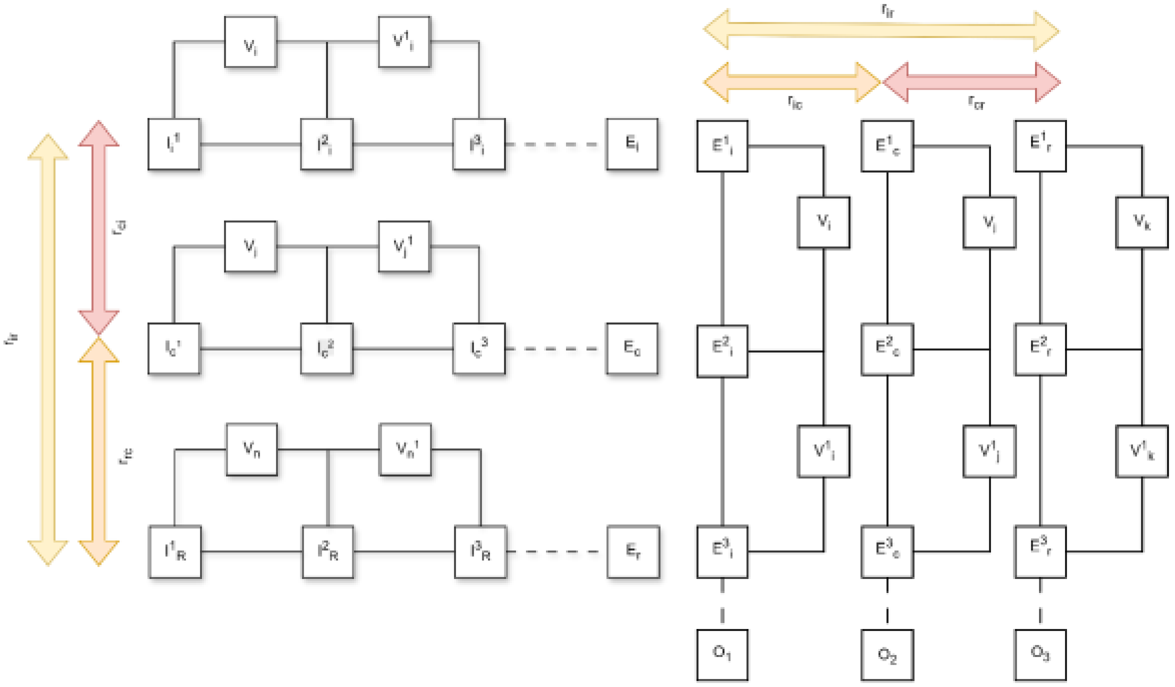

Figure 3: Two layers of the proposed multilayered GNN algorithm.

The values from the cluster level are aggregated to understand the overall behavior and state of the cluster. The aggregation function can be represented using the formula (19).

$$H_c(i) = f(\{h_v(i) | v \in c\}) \quad (19)$$

After concatenating all the outputs from the layer, the final outputs for each variable are obtained using the Multi-Layer Perceptron (MLP) layer as represented by the formula (20).

$$o_v = MLP(concat(h_v(1), .. h_v(i))) \quad (20)$$

The entire work of our proposed multilayered GNN algorithm is outlined in the steps below.

Algorithm: Multilayered GNN algorithm.

Input: Age, health, movement behaviors, location, initial parameters for the SEIR model.
Outputs: Final parameters for the SEIR model.

Step 1: Initialize the individual as a node n_i.
Step 2: Initialize clusters (Eg, Home, workplace, grocery store) as $c_i(i)$.
Step 3: Create edges as connections between nodes and clusters, e_i and weights, as w_e
Step 4: Initialize geographic location as a bigger cluster $c_g(i)$.
Step 5: The Connection between the clusters within a geographic location is initialized as w_c.

```
Step 5: Create an edge with weights for the connection between geographic locations $w_g$.
Step 6: Calculate the weight between the nodes using the function.
    $w_{ij}{}^{t+\Delta t} = w_{ij}{}^{t}.e^{(-d\Delta t)}$
Step 7: Estimate the loss function for predicted vs actual values using the formula
    $L_t = \sum_{i=1}^{n} I_{act_t}(i) - I_{pred_t}(i)$
Step 8: The temporal and spatial representation of each data point is calculated using the proposed multilayered GNN algorithm.
Step 9: The output values are compared to the actual values to find the RMSE results.
Step 10: End
```

Using that function, the value for all the different variables is generated using the GNN algorithm.

$$r_t(i) = MLP(h_i)$$

Data processing

The data for processing is collected from the Ministry of Health and Family Affairs of India [21]. The dataset consists of age, gender, medical condition, and mortality information for patients, which is used to train the demographic model. We also feed real-time case numbers to the model to update error values and continually learn during training. The training and testing of the GNN algorithm are done in a PyTorch environment. The updated results are compared and presented in the section below.

Results and discussion

Considering the infection data from India, we used COVID-19 infection data and recovery data from January 1, 2021, to October 28, 2021. Two case studies analyze the operation of the proposed model in real-world scenarios. The first environment is set where no control measures are taken into consideration, and people's movement within the environment is not controlled. The movement factor $M_h It$ is set to initial conditions and changes according to the formula. However, in the second case study, the movement factors are controlled, and we will examine the impact of these changes on the progression of the pandemic.

The simulation of the case study comprises 500 samples with varying ages and gender groups. The initial condition for SEIR values is assigned by the GNN algorithm using previous training data. The initial population spread is random, and movement and interaction values are assigned based on historical values. With these initial conditions, the progression of the pandemic under two different circumstances is measured and analyzed as shown below.

Case Study 1: Pathogen Circulation in the System.

In this model, the transfer of pathogens between the host and the environment without any control is measured. In the simulation of this model, the host transfers between the environment and interacts with it without any restrictions. In many simulation environments, only one type of pathogen transfer is considered; however, our proposed model takes into account both host-to-target and environment-to-target scenarios before calculating the SEIR values. The equation can best describe the population dynamics in this environment.

$$N_t = N_{t-1} + \Delta N_h + \omega N_e$$

Where$N_{\Delta Th}$is is the total pathogen outflow due to infection caused by host transfer, and ωN_erepresents the total outflow of the population due to the environmental transfer of pathogens in these homogeneous environments. The ecological impact factors also consider the death of pathogens as well as their growth within the environment. For simplicity of modeling, we only regarded the bidirectional pathogen transfer from host to agent and from medium to agent. However, when considering the dynamics of spreading through multiple other ways, it can be exhausting and negate the motive of the modeling. Therefore, to contain the complexity of the model, we will focus on the host-to-agent and medium-to-agent models for the time being. The simulation of the pandemic is shown in Figure 4 below.

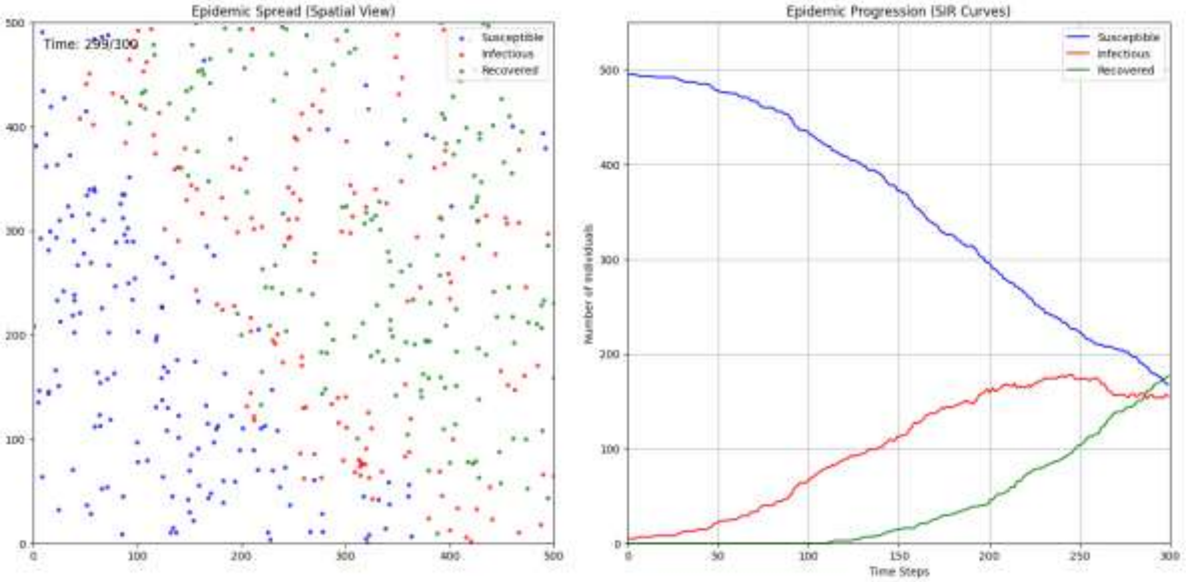

Figure 4: Simulation of the pandemic without any constraints.

There is no intervention to stop the progression of the pandemic, and we can see that the number of infectious people continues to increase exponentially, while the recovery rate also increases. However, in a real-world setting, this increased infection cannot be handled by a limited hospital setup. Therefore, in this environment, the whole healthcare system will collapse, leading to an increased death rate.

Case study 2: Infection control through social distancing.

In the other model scenarios, control measures such as limiting movement and maintaining social distancing between agents are taken into consideration. In these models, we do not consider sanitization, vaccination control methods, or the reduction in infection that occurs due to limiting the movement of people within the cluster. Using this control mechanism alone, the total infection rate of the pandemic came below the allowable limit, as shown in the figure below. Existing hospital settings can control this reduced infection rate. Therefore, a simple act of social distancing, strictly followed by everyone, can reduce the infection rate and save many lives. When this mechanism is combined with other control measures, such as vaccination and mask-wearing, it can further enhance the effectiveness of pandemic control.

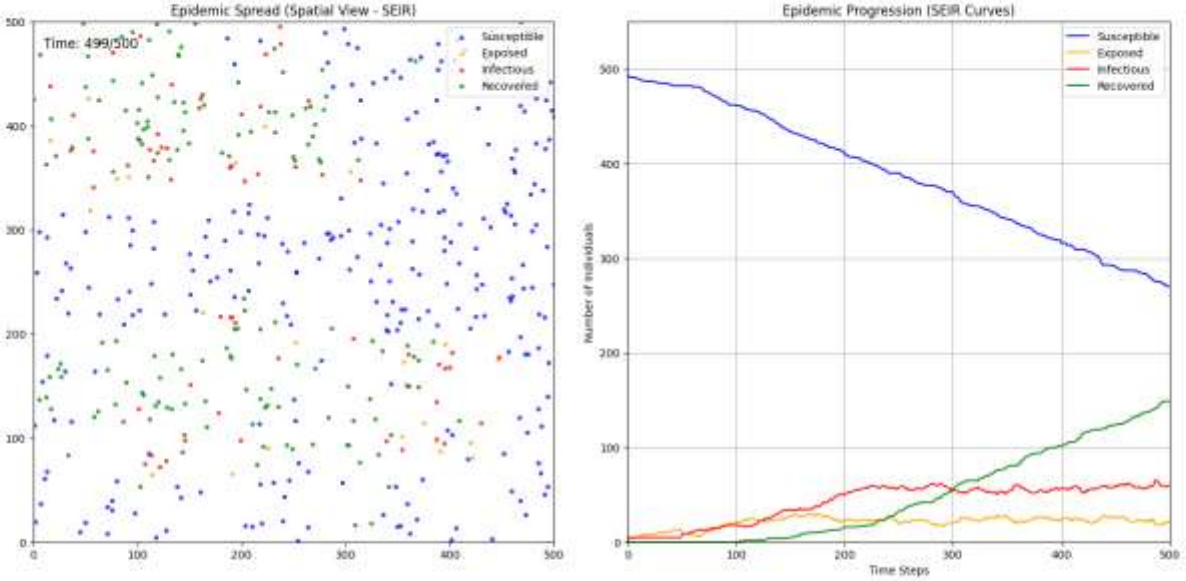

Figure 5: Pandemic model with control measures.

Performance evaluation:

To evaluate the performance of the proposed model, we calculated the MAE and RMSE values between the actual pandemic infection values and the predicted infection values over 100 days. The data was collected from the Indian COVID-19 database and used to verify the performance of the proposed models. The prediction results of infection data are compared to real-time data in Figure 6 below. The predicted data shows a similar trend, similar to real-world data, without any input of real-world information. In some areas, there is a significant deviation from the original value, increasing the number of infectious cases. The Root Mean Square Error (RMSE) value is lower than the SEIR model values, indicating a more powerful predictive performance. The Mean Absolute Error (MAE) value is also lower than the average prediction of the SEIR model.

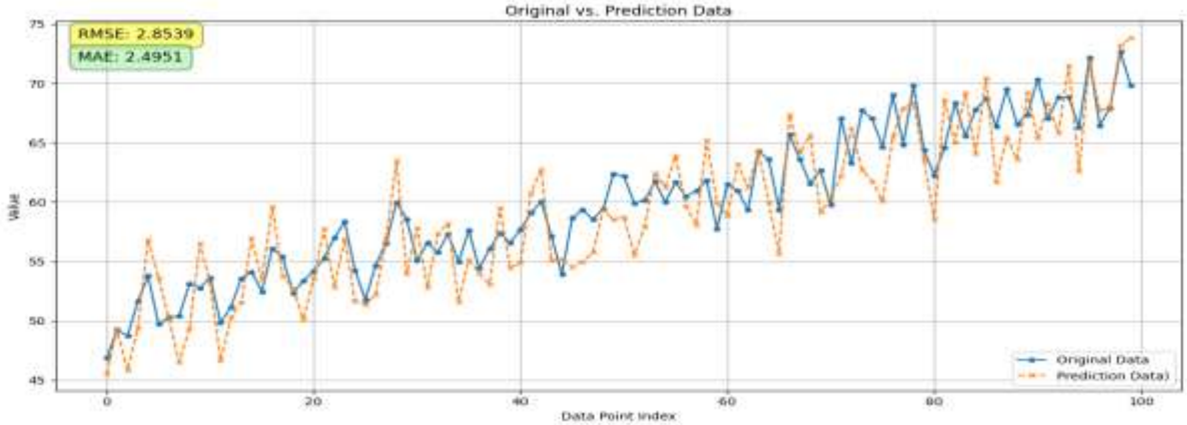

Figure 6: Performance evaluation of the proposed model.

The results indicate that the model's performance is suitable for both long-term and short-term predictions. The loss curve in Figure 7 suggests that the overall loss of the model increases initially and then decreases over time, demonstrating the learning capability of the GNN model. The GNN model's learning and risk scores based on infection show that the risk values are gradually decreasing with every time step, indicating the model's capability to learn new data and predict results more accurately.

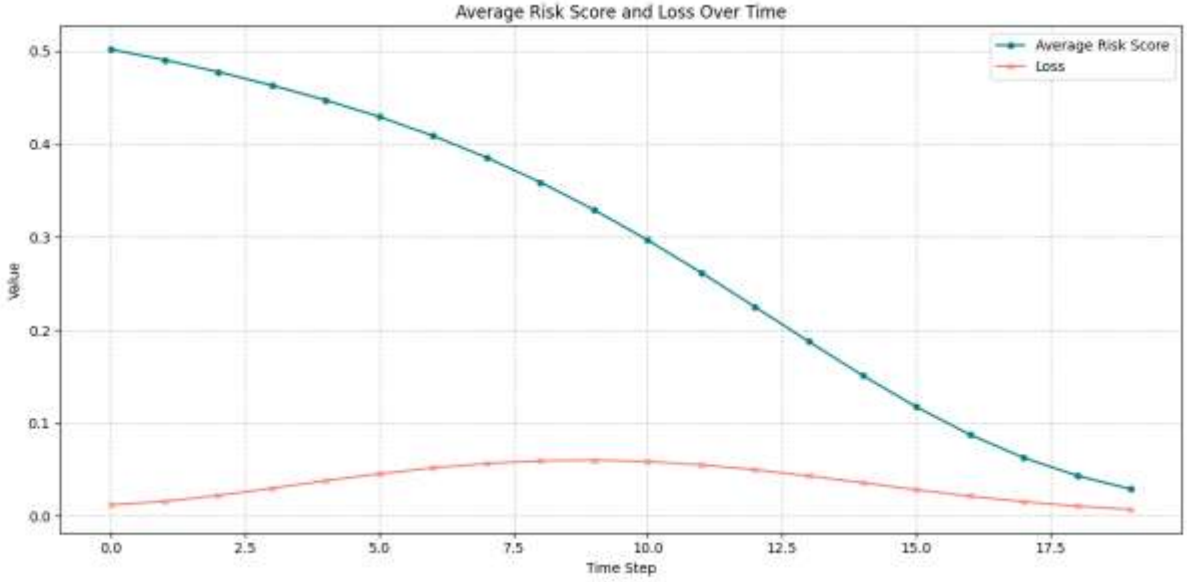

Figure 7: Loss curve of the proposed GNN model.

When comparing the prediction results with actual results by manually adjusting key parameters, such as incubation period and recovery rate, which can be altered according to stringent governmental rules, we found that the predicted results and actual results follow the same curve. The exact values of the results are off by 5 to 10%, but both values follow similar trends, as shown in Table 1 below.

Table 1: Predicted values vs actual values for different input parameters.

Σ	Γ	α	Actual value	Peak	Predicted value
0.5	0.2	0.6	15155	Day 258	15156
0.45	0.18	0.9	13871	Day 232	13729
0.4	0.16	0.6	15221	Day 261	17851
0.35	0.14	0.9	14060	Day 236	16521

The proposed hybrid model, which integrates the GNN algorithm with the SEIR mathematical model, has improved performance compared to existing compartmental models, as shown in the table below. The results show that the total error rate is reduced by up to 75%, and the MAE is reduced by 50%, indicating a significant improvement over previous models.

Table 2: RMSE comparison.

Models	RMSE	MAE
SEIR+GNN	2.85	2.49
SEIRV	10.66	4.8

Conclusion

The proposed hybrid model outperforms existing standalone compartmental models in predicting the pandemic's outcome over a long time interval. The proposed multilayered GNN model does not require a large amount of training data, reducing the need for historical data. The proposed model strategy, which utilizes dynamic risk scoring and clustering techniques, has addressed the limitations of the traditional SEIR model. The proposed model has three times higher accuracy in predicting the pandemic data compared to existing SEIR models. This hybrid approach can be used in epidemiology fields to reduce error rates. The study provides a comprehensive framework for future researchers to integrate mechanistic models with machine learning models, enabling highly scalable and adaptive systems. The proposed model can be used for any novel pandemic outbreak. Therefore, they can be used in any real-world pandemic situation.

Reference

[1]Dietz, K., & Heesterbeek, J. A. P. (2002). Daniel Bernoulli’s epidemiological model revisited. Mathematical biosciences, 180(1-2), 1-21.

[2]Rao, B. L. (2003). Epidemiology and control of influenza. The National Medical Journal of India, 16(3), 143-149.

[3]Ash, C. (2015). Mathematical modeling of infectious diseases. Science, 347, 1213-1213.

[4]Heesterbeek, H., Anderson, R., Andreasen, V., Bansal, S., De, D., , A., Dye, C., Eames, K., Edmunds, J., Frost, S., Funk, S., Hollingsworth, D., House, T., Isham, V., Klepac, P., , J., , L., Lloyd-Smith, J., Metcalf, J., Mollison, D., Pellis, L., Pulliam, J., Roberts, M., Viboud, C., & Newton, I. (2015). Modeling infectious disease dynamics in the complex landscape of global health. Science, 347.

[5]Chen, S., Robinson, P., Janies, D., & Dulin, M. (2020, August). Four challenges are associated with the current mathematical modeling paradigm for infectious diseases, necessitating a shift. In Open forum infectious diseases (Vol. 7, No. 8, p. ofaa333). US: Oxford University Press.

[6]Hamilton, A. J., Strauss, A. T., Martinez, D. A., Hinson, J. S., Levin, S., Lin, G., & Klein, E. Y. (2021). Machine learning and artificial intelligence: applications in healthcare epidemiology. Antimicrobial Stewardship & Healthcare Epidemiology, 1(1), e28.

[7]Ye, Y., Pandey, A., Bawden, C., Sumsuzzman, D. M., Rajput, R., Shoukat, A., ... & Galvani, A. P. (2025). Integrating artificial intelligence with mechanistic epidemiological modeling: a scoping review of opportunities and challenges. Nature Communications, 16(1), 1-18.

[8]Al Meslamani, A. Z., Sobrino, I., & de la Fuente, J. (2024). Machine learning in infectious diseases: potential applications and limitations. Annals of Medicine, 56(1), 2362869.

[9]Fontinele, F., Citeli, M., & Cintra, P. (2020). Mathematical Models for Describing and Predicting the COVID-19 Pandemic Crisis. arXiv: Physics and Society.

[10]Chen, W., Luo, H., Li, J., & Chi, J. (2024). Long-term trend prediction of pandemic combining the compartmental and deep learning models. Scientific Reports, 14.

[11]Guo, Z., Tian, Y., , Y., Chawla, N., & Gerard, P. (2022). Hierarchical Spatio-Temporal Graph Neural Networks for Pandemic Forecasting. Proceedings of the 31st ACM International Conference on Information & Knowledge Management.

[12]Koruthu, L. (2025). Identifying High-Risk Patients in Hospitals to Prevent Complications Using Graph Neural Networks (GNNs) and the COVID-19 Open Research Dataset (CORD-19). INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT.

[13]Hou, Y., & Bidkhori, H. (2024). Multi-feature SEIR model for epidemic analysis and vaccine prioritization. Plos one, 19(3), e0298932.

[14]Soulaimani, S., Kaddar, A., & Rihan, F. A. (2024). Analysis of a fractional endemic SEIR model with vaccination and time delay. The European Physical Journal Special Topics, 1-17.

[15]Ushirobira, R., & Efimov, D. (2020). On an interval prediction of COVID-19 development based on a SEIR epidemic model. Annual Reviews in Control, 51, 477 - 487.

[16]Khastan, A., Long, H., & Dong, N. (2020). Optimal control of a fractional order model for granular SEIR epidemic with uncertainty. Communications in Nonlinear Science & Numerical Simulation, 88, 105312 - 105312.

[17]Wang, J. (2020). Mathematical models for COVID-19: applications, limitations, and potentials. Journal of public health and emergency, 4, 9.

[18]Ahmadini, A. A. H., Raghav, Y. S., Mahnashi, A. M., Islam Rather, K. U., & Ali, I. (2025). Neural networks to model COVID-19 dynamics and allocate healthcare resources. Scientific Reports, 15(1), 15326.

[19]Cheng, C., Aruchunan, E., & Noor Aziz, M. H. (2025). Leveraging dynamics informed neural networks for predictive modeling of COVID-19 spread: a hybrid SEIRV-DNNs approach. Scientific Reports, 15(1), 2043.

[20]Zheng, Y., Li, Z., Xin, J., & Zhou, G. (2020). A spatial-temporal graph based hybrid infectious disease model with application to COVID-19. arXiv preprint arXiv:2010.09077.

[21]Home | Ministry of Health and Family Welfare | GOI. (n.d.). https://www.mohfw.gov.in/