University of Tehran

Earth Observation and Geomatics Engineering

Online ISSN: 2588-4360

Homepage: https://eoge.ut.ac.ir/

Modeling the Propagation of COVID-19 Using a Multilayer Perceptron and Radial Basis Function in Digital Twins Framework

Zeinab Neisani Samani [1, 2], **Ali Asghar Alesheikh** [1, 2 *]

1. Department of GIS, Faculty of Geodesy and Geomatics Engineering, K. N. Toosi University of Technology, Tehran, Iran. E-mail: zeinab.neisani@email.kntu.ac.ir ; alesheikh@kntu.ac.ir (Corresponding author)

2. Geospatial Big Data Computations and Internet of Things (IoT) Lab, K. N. Toosi University of Technology, Tehran 19967-15433, Iran.

Article Info

Article type:
Research Article

Article history:
Received 2024-10-25
Received in revised form 2024-11-27
Accepted 2024-12-06
Available online 2025-01-01

Keywords:
Prediction,
COVID-19,
Spatial Model,
Artificial Neural Network,
Digital Twin.

ABSTRACT

Coronavirus-19 (COVID-19), as the most recent global pandemic, has significantly affected individuals' daily habits and mobility. These alterations have been related to location, and they can be predicted using Geospatial Artificial Intelligence (GeoAI) modeling. GeoAI and geo-visualization serve as essential tools for gaining clearer insights into the application of spatial phenomena in reality. Digital Twin (DT) as a visualization technology combines software and human efforts in research, particularly in healthcare. They create virtual replicas of patients for disease modeling, allowing for personalized medicine by simulating disease progression and treatment responses. This enhances predictive accuracy and helps develop tailored therapeutic strategies. This paper aims to detect spatial patterns and effective criteria in the outbreak of COVID-19 using GeoAI within the framework of DT. The main contribution is the application of kernel-based algorithms to the disease distribution pattern. The applied data is organized into three general categories: infrastructure (distance to road, land use), environment (traffic congestion, air pollution), and socioeconomic **(**population density, gender ratio, income, education level)**.** Each of these categories has its own sub-criteria. The Multilayer Perceptron (MLP) considers the relationships of input targets based on a normal distribution, while the Radial Basis Function (RBF) technique considers the assumption of a radial influence zone. The COVID-19 dataset was collected over four months from eight hospitals in Tehran. The interpretation of the results indicates that the RBF network, with an RMSE of 1.77e-08, models the COVID-19 outbreak more accurately than the MLP, which had an RMSE of 0.0037. The application of DT with MLP and RBF represents a powerful approach to modeling and simulating complex systems. Utilizing the Artificial Neural Network (ANN) algorithm within the digital twin framework, health centers can achieve enhanced predictive capabilities and real-time responsiveness, improving treatment processes across various medical domains.

Cite this article: Neisani Samani, Z., Alesheikh, A.A. (2023). Modeling the Propagation of COVID-19 Using a Multilayer Perceptron and Radial Basis Function in Digital Twins Framework (2023). *Earth Observation and Geomatics Engineering*, Volume 7, Issue 2, Pages 66-77. http//doi.org/ 10.22059/eoge.2024.384353.1161

Publisher: University of Tehran.
DOI: http//doi.org/10.22059/eoge.2024.384353.1161

1. Introduction

The pandemic caused by Coronavirus-19 (COVID-19) has a specific form of outbreak. This is the latest global pandemic. The outbreak of this disease has had a huge impact on changing people's lifestyles and daily movement patterns. Since social distancing was considered the most effective way to prevent transmission, many daily activities such as going to work, shopping, and commuting changed during the outbreak. These changes were location-based and can be predicted using Geospatial Artificial Intelligence (GeoAI) modeling (Baser, 2020; Neisani Samani et al., 2022; Atek et al., 2022).

GeoAI and digital geovisualization are technologies for a finer perception of the actions of spatial phenomena in the real world (Uhlenkamp, 2019) and such a significant capability in virtual space. The term "Digital Twins" (DT) as a type of as a visualization technology in virtual space has generated significant interest in academia and business (Uhlenkamp, 2019; Callcut et al., 2021). DT is a virtual representation of physical assets and interacts with its real-world counterpart (Andrade, 2021). Key capabilities of digital twin technology include real time monitoring, planning, optimization, maintenance, and remote access. As digital twins are applied across different sectors, their application could rise significantly in the upcoming decades (Singh, 2021; Chen et al., 2021; Agrawal et al., 2023). DT predicts the performance of a product or process. DT utilizes advanced analytics capabilities, monitoring, forecasting, testing processes, and service providers will save costs and time in processes (Guo, 2021; Quilodrán-Casas et al., 2022).

In this paper, the pattern of disease outbreaks is predicted using GeoAI methods and the results are shown in a digital twin virtual space. Artificial Neural Network (ANN) algorithms include MLP (Multilayer Perceptron) and RBF (Radial Basis Function) have been used in this research. ANNs methods are adept at regression tasks, making them suitable for predicting future states of the system based on historical data. Algorithms can process and analyze data or process responses to changes in a monitored physical system. MLP could handle the relationships between input and target based on a normal distribution, while RBF considers the assumption of a radial influence zone. The DT framework facilitates data processing and enhances predictive capabilities for various applications. By analyzing data patterns using MLPs and RBFs, unusual actions and potential problems are identified. Also, spatial data can enrich health databases. Comparing the results of each algorithm, a more accurate method can be identified and utilized to diagnose the infectious disease outbreak model. Also, the DT proactive monitoring is crucial for maintaining system integrity and performance (Kamel Boulos et al., 2019; AlZyoud et al., 2022; Jiang et al., 2022).

The latest researches were reviewed to identify the impact of environmental and spatial criteria on the spread of pandemics such as COVID-19 and the application of digital twin technologies and artificial intelligence in the field of geographic information system (GIS) health. These studies can be categorized into several subjects. The first category focuses on research that identifies the critical criteria for the propagation of COVID-19, utilizing artificial neural networks (ANN) and geographic information systems (GIS) methods. Rahnama and Bazargan (2020) identified areas and populations at higher risk for disease in Iran. A GIS serves as a valuable tool for managing and analyzing spatial information. This research focuses on a geospatial study of COVID-19 to model its epidemiological spread within the country. Neysani Samany et al., (2021) studied the highest-risk locations around hospitals treating COVID-19 in Tehran, using the MLP-ANN method. The findings show a meaningful correlation between how far patients are from the hospital and the number of patients present. The sensitivity analysis revealed that patient transaction volumes and the distance to the hospital, along with non-residential hospital usage, are the two key elements influencing the virus's transmission. Neisani Samani et al., (2022) assessed the spatio-temporal traffic flow trends in Tehran before and during the COVID-19 pandemic through non-linear regression analysis. The findings indicate that the time series ANN model developed is capable of accurately forecasting spatio-temporal traffic volumes. Melchane et al., (2024) employed machine learning approaches and feature extraction techniques to improve the forecasting of COVID-19 cases. The techniques used in their research comprised Deep Neural Networks (DNN), Random Forests (RF), and Extreme Gradient Boosting (XGBoost). Results indicate that the deep learning approach outperforms the other two in terms of accuracy.

Another set of studies examined the use of diagnostic tools in disease identification. Allen et al., (2021) developed a DT model for stroke patients that assesses medical characteristics to predict potential pathways leading to acute clinical events. This model goes beyond simply using global illness codes by incorporating laboratory values as inputs. The findings indicate that the DT model can enhance clinical decision-making and provide virtual management resources for more efficient medical trials through accurate predictions of patient pathways. Quilodrán-Casas et al., (2022) introduced an urban model leveraging artificial neural networks and digital twins to analyze COVID-19 in the British. This model investigates spatial changes. Predictions are based on Long Short-Term Memory (LSTM) techniques and compared with predictions based on Generative Adversarial Networks (GANs). Chen et al., (2022) evaluate using artificial intelligence algorithms and digital twins to improve COVID-19 management in smart cities. They identify shortcomings in current public health strategies and introduce a digital twin model called COVID-DT, which combines blockchain technology and deep learning to simulate epidemic prevention and control efforts. The simulation results indicate that the COVID-DT model achieves better results. The model provides a robust framework for enhancing information security and epidemic management in smart cities. Zhang et al., (2023) examined the spatial distribution of the Omicron variant of COVID-19 across the United States. Their research used space-time scanning statistics, the Hoover index, and disease spread trajectories to analyze the spatiotemporal dynamics of the Omicron wave. Neves et al., (2023) highlight the influence of climate on the distribution, incidence, and mortality rates associated with COVID-19. In Brazil, from 2020 to 2021, a modeling approach was employed to evaluate the climatic conditions conducive to COVID-19 cases, focusing on metrics like cumulative incidence, mortality, and fatality rates. To assess the effects of temperature, precipitation, and humidity, seven statistical algorithms were applied. Findings revealed that the annual temperature range and the seasonality of precipitation had a notable impact on case distribution. Higher incidence rates were observed in Brazil's northern and southern regions, whereas the Midwest and Southeast regions had elevated mortality and fatality rates. The results are that although social, viral, and human factors contribute to the spread of the disease, climate serves as a significant co-factor, particularly in regions with high climatic suitability during the analyzed timeframe.

The difference between the present research and previous studies is that this paper focuses on two main aspects: the learning process and the 3D geovisualization of the results in a virtual space, and using kernel-based algorithms to analyze disease distribution patterns. The highlights of the paper are such as: DT technology is used as a simulation environment to illustrate disease outbreak patterns. The MLP and RBF were applied to consider the relationships of input targets based on a normal distribution and the assumption of a radial influence zone. The MLP considers the relationships of input targets based on a normal distribution, while the RBF technique considers the assumption of a radial influence zone.

By utilizing the methods proposed in this paper, we can gain more precise insights into health events and disease outbreaks. This enhanced understanding can bolster the preparedness of healthcare professionals and optimize budget allocation within medical facilities.

1. 1. Study Area

Tehran is the most populous city and capital of Iran, the capital of Tehran Province and Tehran County. The city had a population of 9,039,000 in 1401 and, according to the 2018 United Nations estimate, is the 34th most populous city in the world and the most populous city in West Asia. The Tehran metropolitan area is also the second most populous metropolis in the Middle East. (See Figure 1). This area has many clinics and hospitals that accepted patients during the outbreak of COVID-19. This research uses data from 8 hospitals in Tehran dedicated to treating COVID-19 for a period of 4 months. (Neisani Samani et al., 2023).

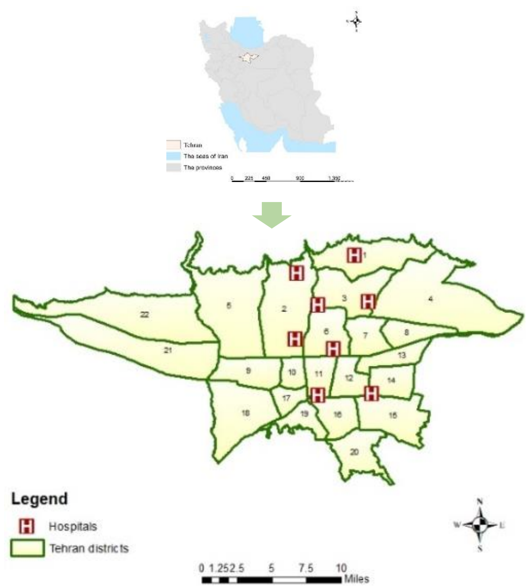


Figure 1. Study Area Overview.

2. Methods and Materials

This paper proposes a four-stage procedure. The first step involves capturing data related to COVID-19 patients and relevant spatial factors. Then, a 3D representation of the data is created. In the third stage, the two algorithms, MLP and RBF, were implemented separately and their results were compared. The final step involves strategic decision-making. All these steps are framed within the digital twin framework (see Figure 2).

The ANN algorithms considered include RBF and MLP. Figure 3 is a flowchart of the research method.

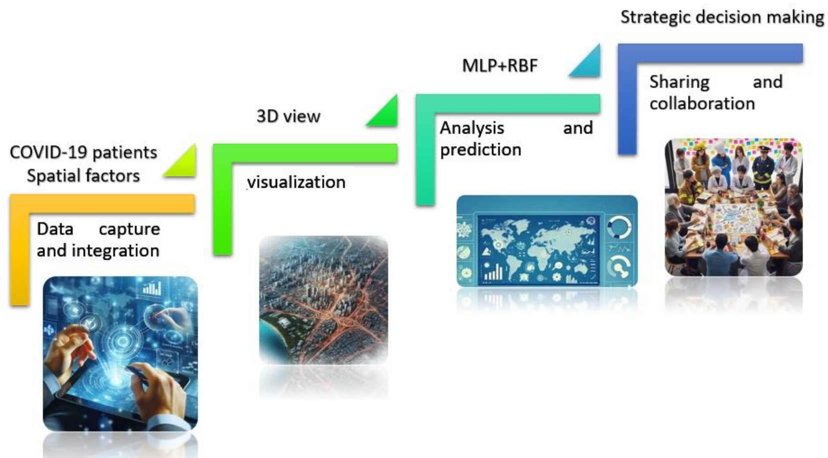


Figure 2. The framework for digital twins and its associated steps**.**

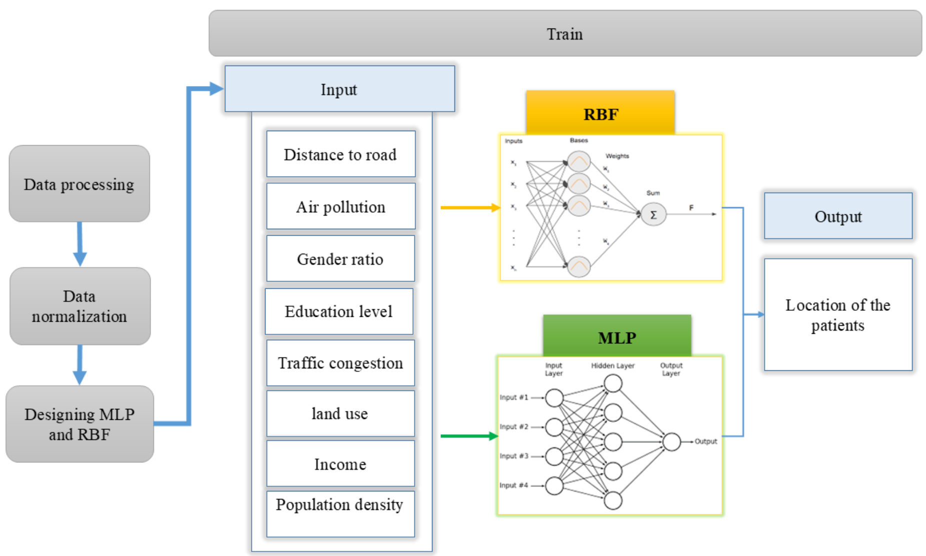


Figure 3. Flowchart of the Proposed Method.

2.1. MLP

An MLP is used in many problems that involve complex nonlinear equations, with the aid of appropriately defined weights and activation functions. Activation functions are used in neurons based on the suitability of the specific task. Neurons are the fundamental components of layers in neural networks. An MLP consists of three types of layers, called the input layer, the hidden layer, and the output layer. Neurons in each layer are interconnected with neurons in the other layers, but they do not connect to other neurons within the same layer (Ghanou & Bencheikh, 2016; Neisani Samani et al., 2020). A key difficulty in these networks is figuring out the ideal quantity of hidden layers and the number of neurons per layer, as there are differing opinions on this matter. It has been demonstrated that an ANN with one hidden layer utilizing a sigmoid activation function as stated in Eq.1, a linear activation function in the output layer can accurately approximate any desired function to any level of precision, assuming there are enough neurons present in the hidden layer (Martins et al., 2004).

$$O = 1/1 + e\ net \qquad (1)$$

In an artificial ANN model, inputs are entered into neurons represented by the vector *x*. Each neuron receives an input signal, with a corresponding communication weight. This weight vector, denoted as *W*, consists of individual weights w_1......, w_n multiplied by each input signal. The resulting values are aggregated within the neuron to compute the output value (Martins et al., 2004) (Eq. 2):

$$\text{NET} = \sum_{i=1}^{n} x_i w_i \quad (2)$$

2.2. RBF

RBF is a type of ANN with three layers: input, intermediate / hidden, and output. The input layer has high dimensionality and contains multiple nodes. The intermediate layer has n nodes, each centered on specific centers Cn. In the intermediate layer, the process begins by calculating the distance from the input vector to the designated center, followed by the applying of the transfer function φ. The output, denoted as y, is then generated as a numerical value in the output layer. The output layer performs a linear transformation that combines the values from the intermediate layer. In this context, N typically represents the number of intermediate layer nodes, while m denotes the number of input data points. The RBF algorithm is fundamentally based on the distance from a given point to the center of the intermediate layer. If the relationship $\varphi(x) = \varphi(||x||)$ holds, φ is classified as a radial basis function. Therefore, φ applied to vectors in R_N space, it is expressed just as a function of distance. The function Φ can take various numerical forms. The points illustrate the implementation of the Euclidean Distance Matrix (EDM) using the Φ function (See Eqs. 3 and 4) (Haijun et al., 2023).

$$EDM = \begin{bmatrix} ||x_1 - x_1|| & ||x_1 - x_2|| & \cdots & ||x_1 - x_n|| \\ \vdots & & \ddots & \vdots \\ ||x_n - x_1|| & ||x_n - x_2|| & \cdots & ||x_n - x_n|| \end{bmatrix} \quad (3)$$

$$f(x) = a_1 \emptyset (||x - x_1||) + \cdots + a_1 \emptyset (||x - x_n||) \quad (4)$$

The function $f(x)$ estimates the constant coefficients of a function represented by $[a_1, \ldots\ldots, a_n]$. The symbol Φ denotes the transfer function, which is a nonlinear function mapping from $R^+ \to R$. Some radial basis transfer φ depends on the distance function and a shape coefficient (ε); they also take into account the increase in length and the expansion of the function. It is the standard deviation of the normal distribution (σ), where $\varepsilon 2 = 1/2\sigma 2$ (Fornberg & Piret, 2008). The transformation of the elements is $\Phi(A)_{ij} = \Phi(A_{ij})$, it is used to convert the numerical function φ into the matrix φ. In the context of the EDM, the function φ is utilized, and the resulting matrix is referred to as the transfer matrix. The influence of the transfer function on the radial estimate can be evaluated. Once the type of transfer function is specified, the coefficients $a_1, \ldots, a_n$ in the input data $x(x_1, x_2, \ldots\ldots, x_n)$ can be determined by solving the p-number (See Eq. 5) (Haijun et al., 2023):

$$\begin{gathered} a_1 \emptyset (||x_1 - x_1||) + \cdots + a_p \emptyset(||x_1 - x_n||) = y_1 \\ a_1 \emptyset (||x_2 - x_1||) + \cdots + a_p \emptyset(||x_2 - x_n||) = y_2 \\ \ldots\ldots \\ a_1 \emptyset (\| x_p - x_1\|) + \cdots + a_p \emptyset(||x_n - x_n||) = y_m \end{gathered} \quad (5)$$

This set is valid if the vectors $x_i \neq x_j$ for $i \neq j$ and the $n \times n$ matrix derived from the set of equations is inverted. All p data points in this model can utilize to balance complexity and accuracy; alternatively, the number of k points $C_1, \ldots, C_k$ can consider as the centers for the Radial Basis Function, where $k << n$ (Eq. 6) (Haijun et al., 2023):

$$f_p(x) = a_1 \emptyset (||x - x_1||) + \cdots + a_p \emptyset(||x - x_n||) \quad (6)$$

$$f_k(x) = a_1 \emptyset (||x - x_1||) + \cdots + a_k \emptyset(||x - x_n||)$$

Assuming the output is multi-dimensional, the function is generalized; specifically, it involves the vector of coefficients, denoted as α, and the matrix of coefficients, denoted as W. The RBF network differs slightly from the display mode, as it consists of two mapping layers. First, the dimensions of the mapping space transform from Rn to R_K, and then to the output layer Rm (See Eq. 7). Here, k is the number of rows for the kernels, and n is the number of targets:

$$x \to \begin{bmatrix} ||x - C_1|| \\ ||x - C_2|| \\ \cdot \\ \cdot \\ \cdot \\ ||x - C_k|| \end{bmatrix} \to \emptyset \begin{bmatrix} ||x - C_1|| \\ ||x - C_2|| \\ \cdot \\ \cdot \\ \cdot \\ ||x - C_k|| \end{bmatrix} \to W \begin{bmatrix} \emptyset (||x - C_1||) \\ \emptyset (||x - C_2||) \\ \cdot \\ \cdot \\ \cdot \\ \emptyset(||x - C_k||) \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{bmatrix}$$

$$y_h(x) = \sum_{j=1}^{k} W_{jh} \; \emptyset_j (\|x - C_j\|) \quad (7)$$

In Eq. 7, W represents the matrix of output weighting coefficients from the hidden layer to the output layer (W_{jh}). The dimension of W is increased by maintaining the linearity of the problem and adding the bias vector b. The network train with fixed centers (points) $C_1, \ldots, C_k$, using the data to determine the weight matrix and the bias vector b. This process can be carried out with all the data once, or by updating the weight function in several steps while increasing the number of RBF kernels.

2.2.1. Network training of the RBF

To perform training operations, the information is split into training and validation sets. Then, the wide variety and proximity of RBF facilities and the characteristics of switch φ are determined. Training by way of performing the problem of linear algebra for weights and biases, a system of equations is shaped using every enter pair x and every output y. Here, W represents an $m \times k$ weight matrix, Y denotes the $m \times p$ matrix formed by p column vectors in IRm. The matrix representation of the equations to solve is outlined. To address this collection of matrix equations, refer to Eq. 8:

$$W\Phi = Y \tag{8}$$

Where Φ is a $k \times p$ Transpose Matrix and each column represents the function of the corresponding base radius (See Eq. 9).

$$\Phi_j = \varphi_j (\| x - c_j \|) \tag{9}$$

This should increase the dimension of the matrix Φ by adding another row for the bias section. To optimize the weight matrix, an appropriate error function is employed. A standard method is the sum of squares error function (See Eq. 10).

$$E = \frac{1}{2} \sum_{i=1}^{n} \sum_{h=1}^{m} [y_{h\ (x_i) -\ t_{h,i}}] \text{^}2 \tag{10}$$

Which $t_{m,n}$ is the target value for the network output y_m, when the enter vector of the network is x_n. The matrix W is calculated by inverting Φ (See Eq. 11):

$$W = Y\Phi^{+} \tag{11}$$

If the determinant of the matrix Φ is zero, value analysis must be used to solve the equation Φ^{-1}. Using the results of the weight and bias matrices, along with the RBF network matrix, the value of the variable at a new point can be determined.

3. Implementation

In this section, the research implementation method is explained. The first part introduces the applied data organized into three general categories: infrastructure, environment, and Socioeconomic**.** Each of these categories has its own sub-criteria (Neisani Samani & Alesheikh, 2019; Foruzandeh et al., 2024). Table 1 presents the criteria and sub-criteria used in the research. The second part describes the implementation of ANN algorithms, detailing the algorithmic processes and results obtained from the research. The third part outlines the sensitivity analysis.

3.1. Data

The proposed algorithm is run to evaluate the influence of infrastructural, environmental, and socioeconomic factors in the study area. The applied data include a normalized map of the criteria, such as distance to road, land use, traffic congestion, air pollution, population density, gender ratio, income, and education level (See Table 1), which serve as input parameters for the designed ANN-MLP and ANN-RBF models, along with the location of patients over four months from eight hospitals in Tehran in 2020.

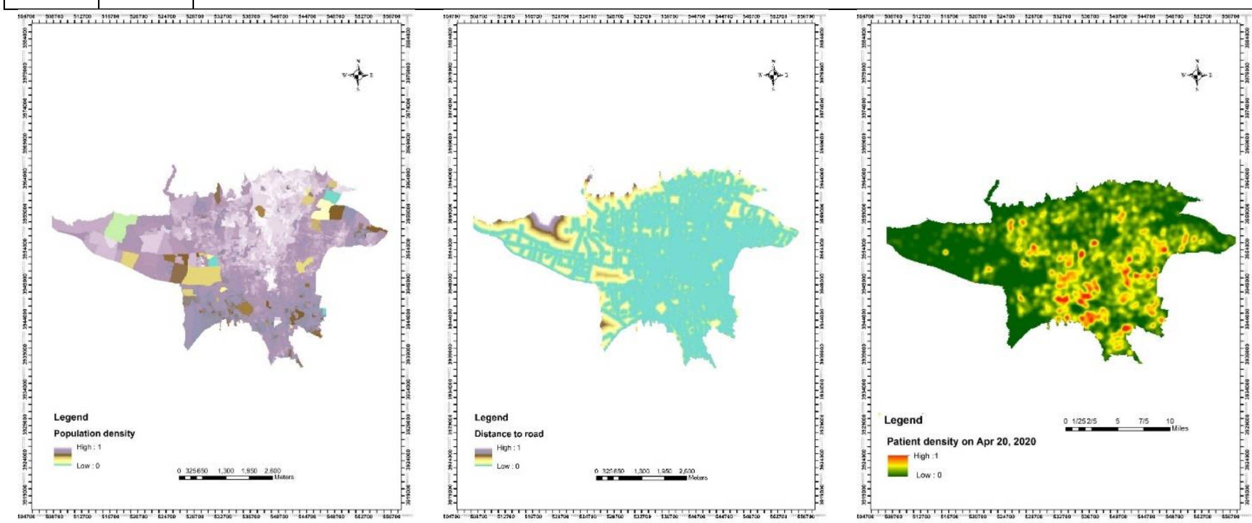
Figure 4 displays the effective criteria and the criteria with less impact. The results of implementing the method in this paper indicate that the two criteria of land use and population density have the most impact, and the criterion of distance from the road has the least impact.

Figure 5 displays the output of the disease prediction model implemented within the framework of the DT.

Table 1. Criteria and sub-Criteria applied in research.

Infrastructural	**Distance to road**	A road is a route on the ground between two places, made so that it is possible to travel between those two points. The road density ratio depends on the road type (Flügel et al., 2022).
	Land use	Land use involves managing and modifying the natural environment to create environments such as settlements and semi-natural habitats (Kartawisastra et al., 2022).
Environmental	**air pollution**	Air pollution is one of the leading environmental causes of diseases and premature deaths worldwide. Pollution sources emit a mixture of all pollutants (Valk et al., 2021;Yates et al., 2022).
	Traffic congestion	The patterns of urban road traffic congestion during the COVID-19 pandemic were markedly different from those observed in the days leading up to it (Almatar, 2023).

Socioeconomic	**Education level**	The number of years of education a person spends. (Stroffolini et al., 2020).
	Population density	Population density is a way of measuring the population per unit area or volume that has been used many times for living organisms, especially humans. The unit of population density is people per square kilometer. Population density is one of the most essential characteristics of an urban area (Baser, 2021).
	Income	In economic terms, income refers to the compensation received for utilizing production factors that individuals possess, which can manifest as wages, rental income, interest, or profits (Orbawati, 2022).
	Gender ratio	The gender ratio within a population is defined as the comparative number of individuals of one gender to that of the opposite gender (Grech, 2015).



(a) (b) (c)

Figure 4. The map illustrates the criteria used in the study. The two criteria include: population density and land use, have the greatest impact on disease prevalence, but the criterion of distance from roads has the least impact. The components are: (a) Population density, (b) Distance to roads, (c) Patient density.

Figure 5 illustrates a segment of the study area. Shariati General Hospital is located in District 6 of Tehran Province, on Jalal-e-Ale Ahmad Street. It served as a critical medical center during the COVID-19 outbreak, treating many patients. The red circles represent the pattern of disease outbreaks, with Shariati Hospital at the center of the largest circle. The surrounding buildings consist of residential and commercial properties.

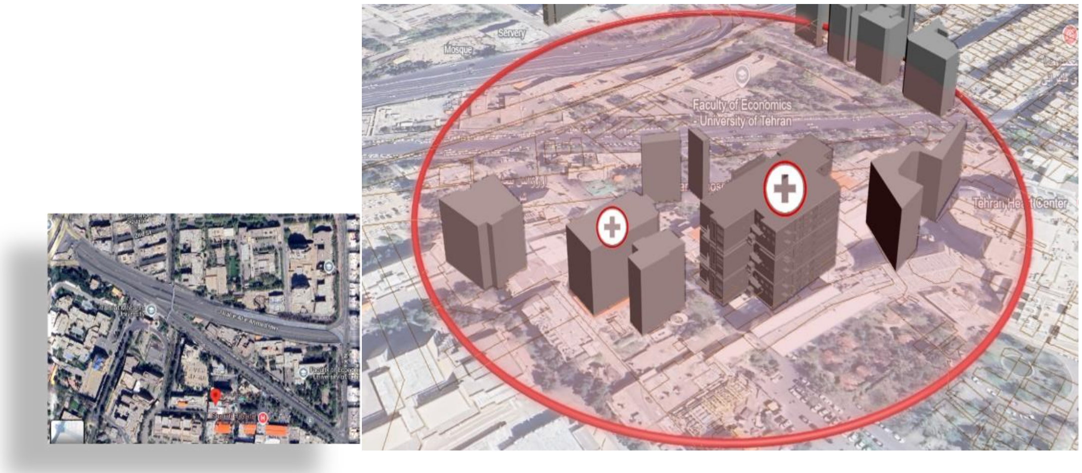
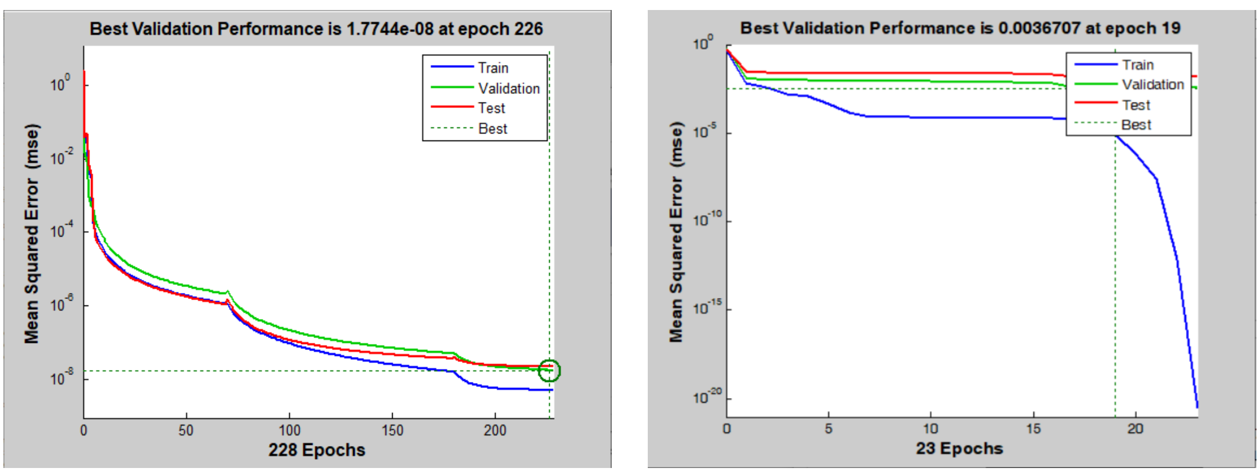


Figure 5. Output of the Disease Prediction Model Implemented within the Digital Twin Framework (Right figure) (The location of the hospital is plotted on a satellite image that is the base map of Google Maps (Left figure)).

3.2. Running MLP and RBF

Figure 6 presents the results of running MLP and RBF. The interpretation of the results indicates that the RBF network, with an RMSE of 1.77e-08, models the COVID-19 outbreak more accurately than the MLP, which has an RMSE of 0.0037. The architecture of the designed ANN is configured with eight neurons in the input layer, three hidden layers, each containing 15 neurons, and five neurons in the output layer. The activation function used for the MLP is sigmoid, while the RBF employs a quadratic activation function.

The implementation of two applied algorithms in this research is in MATLAB 2016 software. The computer running the above program has the following specifications: Intel Core i7 processor @ 3.20 GHz and 4 GB RAM.



(a) (b)

Figure 6. Performance of a) RBF, b) MLP

3.3. Sensitivity Analysis

Sensitivity analysis in GIS is a vital tool for understanding how variations in input data affect spatial model outputs. This approach is essential in environmental modeling, urban planning, and Health resource management, where decisions based on GIS analysis can have significant implications (Crosetto & Tarantola, 2001).

Variance-Based Techniques: These methods assess how variations in input parameters contribute to the general uncertainty in model outputs. Techniques like the extended Fourier Amplitude Sensitivity Test (FAST) allow for efficient estimation of sensitivity indices without requiring separate model runs for each input variable (Mowbray, 2022).

Monte Carlo Simulations: This probabilistic method involves running simulations multiple times with randomly varied inputs to observe the range of possible outcomes. It effectively captures the uncertainty inherent in complex models (Mowbray, 2022).

One-at-a-Time (OAT) Analysis: This straightforward approach varies one input at a time while keeping others constant, making it easier to identify the impact of individual variables on output results (Mowbray, 2022).

In this paper, OAT was used for spatial analysis, sensitivity analysis was performed by removing each criterion, and the accuracy of the desired artificial neural network was calculated. As a result of the analysis, population density and land use were identified as the most important criteria in the spatial spread of COVID-19. The distance to the road criterion has less impact on the study results. The sensitivity results in MLP and RBF are the same. Table 2 shows the results of the research sensitivity analysis in MLP and RBF.

Table 2. The sensitivity analysis results in RBF and MLP.

Criteria	R in RBF	R in MLP	RMSE in RBF	RMSE in MLP
Distance to road	99.681	99.214	1.77e-5	1.25e-5
Air pollution	99.574	99.415	1/45E-04	1/45E-04
Gender ratio	99.524	99.214	1.42e-5	1.21e-5
Education level	99.414	99.025	1.21e-5	1.58e-5
Traffic congestion	99.258	99.155	1.2e-6	1.14e-6
Landuse	94.247	93.122	1.4e-3	1.27e-3
Population density	93.132	93.541	1.76e-3	1.5e-3

4. Discussions

A DT is a virtual model that reflects a physical system, process, or product, facilitating simulation and analysis. It incorporates real time information to accurately represent the condition and actions of its physical equivalent. In this research, MLP and RBF are forms of ANN utilized within the DT framework to improve predictive functionalities via ANN techniques. The advantages of implementing these approaches in the context of DT are outlined below:

- MLPs and RBFs can effectively handle large datasets, making them suitable for complex DT virtual space applications involving multiple variables and interactions.

- Digital twin technology can be designed for various applications, including industrial processes and healthcare, and can define broad frameworks with specific objectives.

- MLPs and RBFs can enhance the accuracy of predictions made by DT, leading to improved insights and operational efficiencies.

According to the benefits, there are some challenges as stated:

- Training MLPs and RBFs requires significant computational resources, significantly as the complexity of the model increases with more layers and neurons.

- There is a potential for overfitting when using MLPs and RBFs, particularly with limited training data. Regularization techniques are essential to mitigate this risk.

5. Conclusion

This paper aims to identify the spatial patterns of COVID-19 in Tehran, the capital of Iran. It seeks to determine the most relevant spatial criteria associated with the outbreak of the infectious disease. The methods employed in this study include MLP and RBF networks. The target criteria encompass distance to roads, land use, traffic congestion, air pollution, population density, gender ratio, income, and education level, with the patients' locations serving as the output. The implementation results indicate that the RBF network effectively captures the kernel-based behavior of COVID-19, exhibiting a lower root-mean-square-error (RMSE) and enabling more accurate modeling of the outbreak compared to the MLP network. Digital twins are used to simulate and visualize the dynamics of the disease outbreak. Sensitivity analysis results reveal that population density and land use are recognized as the most important factors in the spatial outbreak of COVID-19, while distance to roads is considered the least influential factor. These sensitivity results are consistent for both MLP and RBF networks.

6. Suggestions

To enhance the study, the following recommendations are provided for researchers: (1) Integrating epidemiological data with real time mobility statistics within the digital twin model can facilitate new investigations, improving the precision and immediacy of predictions related to COVID-19 transmission. (2) The digital twin framework can be utilized to simulate various intervention approaches (such as lockdowns, mask mandates, and vaccination initiatives) and evaluate their influence on the spread of COVID-19. (3) To broaden the scope and relevance of spatial health research, the model established in this study can be applied to understand the transmission dynamics of other infectious diseases.

Conflict of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Acknowledgments and Declarations

Not Applicable.

References

Agrawal, A., Thiel, R., Jain, P., Singh, V., & Fischer, M. (2023). Digital Twin: Where do humans fit in?. *Automation in Construction*, *148*, 104749. https://doi.org/10.1016/j.autcon.2023.104749

Allen, A., Siefkas, A., Pellegrini, E., Burdick, H., Barnes, G., Calvert, J., ... & Das, R. (2021). A digital twins machine learning model for forecasting disease progression in stroke patients. *Applied Sciences*, *11*(12), 5576. https://doi.org/10.3390/app11125576

Almatar, K. M. (2023). Traffic congestion patterns in the urban road network:(Dammam metropolitan area). *Ain Shams engineering journal*, *14*(3), 101886. https://doi.org/10.1016/j.asej.2022.101886

Al-Zyoud, I., Laamarti, F., Ma, X., Tobón, D., & El Saddik, A. (2022). Towards a machine learning-based digital twin for non-invasive human bio-signal fusion. *Sensors*, *22*(24), 9747. https://doi.org/10.3390/s22249747

Andrade, M. A. N., Lepikson, H. A., & Machado, C. A. T. (2021). A new framework and methodology for digital twin development. In *2021 14th IEEE International Conference on Industry Applications (INDUSCON)* (pp. 134-138). IEEE. *https://doi.org/*10.1109/INDUSCON51756.2021.9529701

Atek, S., Pesaresi, C., Eugeni, M., De Vito, C., Cardinale, V., Mecella, M., ... & Gaudenzi, P. (2022). A Geospatial Artificial Intelligence and satellite-based earth observation cognitive system in response to COVID-19. *Acta Astronautica*, *197*, 323-335. https://doi.org/10.1016/j.actaastro.2022.05.013

Baser, O. (2021). Population density index and its use for distribution of Covid-19: A case study using Turkish data. *Health Policy*, *125*(2), 148-154. https://doi.org/10.1016/j.healthpol.2020.10.003

Callcut, M., Cerceau Agliozzo, J. P., Varga, L., & McMillan, L. (2021). Digital twins in civil infrastructure systems. *Sustainability*, *13*(20), 11549. https://doi.org/10.3390/su132011549

Chen, D., AlNajem, N. A., & Shorfuzzaman, M. (2022). Digital twins to fight against COVID-19 pandemic. *Internet of Things and Cyber-Physical Systems*, *2*, 70-81. https://doi.org/10.1016/j.iotcps.2022.05.003

Chen, L., Xie, X., Lu, Q., Parlikad, A. K., Pitt, M., & Yang, J. (2021). Gemini principles-based digital twin maturity model for asset management. *Sustainability*, *13*(15), 8224. https://doi.org/10.3390/su13158224

Crosetto, M., & Tarantola, S. (2001). Uncertainty and sensitivity analysis: tools for GIS-based model implementation. *International Journal of Geographical Information Science*, *15*(5), 415-437. https://doi.org/10.1080/13658810110053125

Flügel, S., Halse, A. H., Hartveit, K. J., & Ukkonen, A. (2022). Value of travel time by road type. *European transport research review*, *14*(1), 35. https://doi.org/10.1186/s12544-022-00554-1

Fornberg, B., & Piret, C. (2008). On choosing a radial basis function and a shape parameter when solving a convective PDE on a sphere. *Journal of Computational Physics*, *227*(5), 2758-2780. https://doi.org/10.1016/j.jcp.2007.11.016

Foruzandeh, M., Neysani Samany, N., & Khodakaramian, B. (2024). A Machine Learning Approach for Modeling the Spatial-temporal Propagation Pattern of COVID-19. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 183-190. https://doi.org/10.5194/isprs-archives-XLVIII-4-2024-183-2024

Ghanou, Y., & Bencheikh, G. (2016). Architecture optimization and training for the multilayer perceptron using ant system. *International journal of computer science*, *43*(1), 10.

Grech, V. (2015). The male to female ratio at birth, Early Human Development,Volume 91, Issue 12,Pages 793-794. https://doi.org/ 10.1016/j.earlhumdev.2015.10.003

Guo, J. (2021). Digital twins are shaping future virtual worlds. *Service Oriented Computing and Applications*, *15*, 93-95. https://doi.org/10.1007/s11761-021-00321-5

Haijun, L., Dai, H., Mao, Y., Wang, L. (2023). An optimized radial basis function neural network with modulation-window activation function. *Soft Computing 28(5):1-18*. https://doi.org/10.1007/s00500-023-09207-4.

Jiang, Y., Li, M., Li, M., Liu, X., Zhong, R. Y., Pan, W., & Huang, G. Q. (2022). Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly. *Automation in Construction*, *141*, 104397. https://doi.org/10.1016/j.autcon.2022.104397

Kamel Boulos, M. N., Peng, G., & VoPham, T. (2019). An overview of GeoAI applications in health and healthcare. *International journal of health geographics*, *18*(1), 7. https://doi.org/10.1186/s12942-019-0171-2

Kartawisastra, Sukarman & Anda, Markus & Suryani, Erna. (2022). Land Use and Land Use Change Strategies. Printed in Indonesia. Indonesian Center for Agricultural Land Resource Research and Development ISBN: 978-602-6916-594.

Martins, A., Neto, A. D. D., & de Melo, J. D. A Neural Network Based Algorithm For Complex Pattern Classification Problems. https://doi.org/10.21528/LNLM-vol2-no2-art4

Melchane, S., Elmir, Y., & Kacimi, F. (2024). Infectious diseases prediction based on machine learning: the impact of data reduction using feature extraction techniques. *Procedia Computer Science*, *239*, 675-683. https://doi.org/10.1016/j.procs.2024.06.223

Mowbray, F. I., Manlongat, D., & Shukla, M. (2022). Sensitivity analysis: a method to promote certainty and transparency in nursing and health research. *Canadian Journal of Nursing Research*, *54*(4), 371-376. https://doi.org/10.1177/08445621221107108

Neisani Samani, Z., & Alesheikh, A. A. (2019). Uncertainty modelling of citizen-centered group decision making using fuzzy-vikor case study: Site selection of healthcare services. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *42*, 809-814. https://doi.org/10.5194/isprs-archives-XLII-4-W18-809-2019

Neisani Samani, Z., Alesheikh, A. A., & Zandi, I. (2023). Development of an approach to creating a spatial equity map of urban health and validation through Volunteered Geographic information (case study: District 6 of Tehran metropolis). *Journal of Geomatics Science and Technology*, *12*(4), 91-106. https://doi.org/10.61186/jgst.12.4.91

Neisani Samani, Z., Alesheikh, A. A., Neysani Samany, N., & Bayat, S. (2022). Assessment of spatiotemporal traffic flow patterns before and during the COVID-19 pandemic using non-linear auto-regressive with external input in Tehran. *Earth Observation and Geomatics Engineering*, *6*(1). https://doi.org/10.22059/eoge.2022.344727.1117

Neisani Samani, Z., Karimi, M., & Alesheikh, A. (2020). Environmental and infrastructural effects on respiratory disease exacerbation: a LBSN and ANN-based spatio-temporal modelling. *Environmental monitoring and assessment*, *192*, 1-17. https://doi.org/10.1007/s10661-019-7987-x

Neves, J. M. M., Belo, V. S., Catita, C. M. S., de Oliveira, B. F. A., & Horta, M. A. P. (2023). Modeling the climatic suitability of COVID-19 cases in Brazil. *Tropical Medicine and Infectious Disease*, *8*(4), 198. https://doi.org/10.3390/tropicalmed8040198

Neysani Samany, N., Toomanian, A., Maher, A., Hanani, K., & Zali, A. R. (2021). The most places at risk surrounding the COVID-19 treatment hospitals in an urban environment-case study: Tehran city. *Land use policy*, *109*, 105725. https://doi.org/10.1016/j.landusepol.2021.105725

Orbawati, E. B., Jalungono, G., Muhammad, I., & Ahsani, R. D. P. (2022). Determinant Income of Merchants in Magelang City During the COVID-19 Pandemic. *KnE Social Sciences*, 167-182. https://doi.org/10.18502/kss.v7i10.11219

Quilodrán-Casas, C., Silva, V. L., Arcucci, R., Heaney, C. E., Guo, Y., & Pain, C. C. (2022). Digital twins based on bidirectional LSTM and GAN for modelling the COVID-19 pandemic. *Neurocomputing*, *470*, 11-28. *DOI: 10.1016/j.neucom.2021.10.043*

Rahnama, M. R., & Bazargan, M. (2020). Analysis of spatio-temporal patterns of Covid-19 virus pandemic and its hazards in Iran. *Environmental Management Hazards*, *7*(2), 113-127. https://doi.org/10.22059/jhsci.2020.304976.571

Singh, M., Fuenmayor, E., Hinchy, E. P., Qiao, Y., Murray, N., & Devine, D. (2021). Digital twin: Origin to future. *Applied System Innovation*, *4*(2), 36. https://doi.org/10.3390/asi4020036

Stroffolini, T., Sagnelli, E., Sagnelli, C., Morisco, F., Babudieri, S., Furlan, C., ... & Almasio, P. L. (2020). The association between education level and chronic liver disease of any etiology. *European Journal of Internal Medicine*, *75*, 55-59. https://doi.org/10.1016/j.ejim.2020.01.008

Uhlenkamp, J. F., Hribernik, K., Wellsandt, S., & Thoben, K. D. (2019, June). Digital Twin Applications: A first systemization of their dimensions. In *2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)* (pp. 1-8). IEEE. https://doi.org/10.1109/ICE.2019.8792579

Valk, J. P., & CCM, J. (2021). The interplay between air pollution and coronavirus disease (COVID-19). *Journal of Occupational and Environmental Medicine*, *63*(3), e163-e167. https://doi.org/10.1097/JOM.0000000000002143

Yates, E. F., Zhang, K., Naus, A., Forbes, C., Wu, X., & Dey, T. (2022). A review on the biological, epidemiological, and statistical relevance of COVID-19 paired with air pollution. *Environmental advances*, *8*, 100250. https://doi.org/10.1016/j.envadv.2022.100250

Zhang, S., Liu, L., Meng, Q., Zhang, Y., Yang, H., & Xu, G. (2023). Spatiotemporal patterns of the Omicron wave of COVID-19 in the United States. *Tropical Medicine and Infectious Disease*, *8*(7), 349. https://doi.org/10.3390/tropicalmed8070349