INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

A PROPOSED DEEP LEARNING FRAMEWORK BASED ON GIS TO PREDICT SPATIAL DISTRIBUTION OF EPIDEMIC INFECTIOUS DISEASES

Ismail S. Tawfik[1], Prof. Christina Albert [2]

[1]*PhD Candidate*, Computer and Information Systems Department, Sadat Academy for Management Sciences, Cairo Egypt, *Ismail.tawfik@gmail.com*
[2] *Prof.* of Computer and Information Systems Department, Sadat Academy for Management Sciences, Cairo Egypt, *sams.christina.albert@gmail.com*

Abstract

Communicable diseases pose significant threats at local, regional, and global levels, often leading to epidemics or pandemics. An epidemic refers to a sudden increase in the number of cases of an infectious disease above what is normally expected in a given population. Examples include cholera, measles, malaria, and dengue fever. Pandemics, however, can result in widespread illness, significant loss of life, and severe social and economic consequences. Concerns about potential pandemic diseases, such as new strains of influenza and severe acute respiratory syndrome (SARS) remain critical.
This study presents a deep learning framework based on Geographic Information Systems (GIS) to predict the spatial distribution of epidemic infectious diseases. The framework combines the strengths of deep learning and GIS techniques, both of which offer exceptional capabilities in the field of epidemiology. The study outlines the key steps involved in developing the proposed framework and explains its operational functionality.
The proposed framework aims to enhance decision-making efficiency, assist governmental authorities in generating sustainable strategies, and establish appropriate protocols to control epidemics, particularly in high-risk areas. By predicting vulnerable areas, the framework helps mitigate the risks associated with outbreaks and protects social and economic stability.
Choosing an appropriate framework requires consideration of several key factors, including those relevant to the spread of diseases or epidemics, accuracy, flexibility, and validation.

Keywords*:* The Analytic Hierarchy Process (AHP), multi-criteria decision analysis (MCDA), Geospatial Artificial Intelligence (GeoAI), Deep Learning (DL), Geographical Information System (GIS), Structure Query Language (SQL), Multi-Layer Perceptron (MLPs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Deep Neural Network (DNN), Hopfield Neural Networks (HNN), Feedforward Neural Networks (FNN), gradient descent (SGD)

1. Introduction

Deep learning is a powerful subset of machine learning that uses artificial neural networks with multiple layers to learn complex patterns from large datasets. This approach has demonstrated exceptional performance in tasks such as image recognition, natural language processing, and predictive modeling, making it highly effective for analyzing diverse and high-dimensional data (Goodfellow, et al., 2016). In the context of epidemic infectious diseases, deep learning models, such as Convolutional Neural Networks (CNNs) and

Recurrent Neural Networks (RNNs), are increasingly used to analyze medical imaging, genomic data, and epidemiological information. These models help predict disease outbreaks, identify high-risk areas, and track the spread of diseases, providing timely insights for public health interventions (Chung, et al., 2020).
Geographic Information Systems (GIS) are integral tools for managing, analyzing, and visualizing spatial data. In epidemiology, GIS is used to map disease outbreaks, identify geographic patterns, and understand the relationship between residential, demographic, and environmental factors that influence the spread of infectious diseases (Ruktanonchai, et al., 2020). GIS allows for the integration of various data sources, such as population density, climate conditions, and infrastructure, to provide critical spatial insights that support disease surveillance and response strategies.
When combined, deep learning and GIS offer a robust framework for epidemic infectious disease management. Deep learning can process large, complex datasets to identify hidden patterns, while GIS provides the spatial context necessary to visualize and predict disease spread. The integration of these technologies enables more accurate predictions of disease outbreaks, better identification of at-risk regions, and more effective resource allocation for controlling the spread of infectious diseases (Zhou, et al., 2021). This synergy is particularly important in real-time epidemic forecasting, early detection, and optimizing public health responses.

2. GIS Concept and Definition

Geographic Information Systems (GIS) are computer-based systems designed for capturing, storing, analyzing, managing, and visualizing spatial and geographic data. GIS integrates hardware, software, and data to enable the manipulation and analysis of geographic information, providing insights into patterns, trends, and relationships in the real world. It allows users to map, model, and interpret spatial data to better understand complex phenomena and make informed decisions in various fields such as urban planning, environmental monitoring, public health, and disaster management (Zhang, et al., 2021).
GIS systems enable decision-makers to visualize geographic data and analyze it in ways that support informed, data-driven decisions. It has become an indispensable tool for spatial analysis, offering solutions for issues ranging from land use planning to disease outbreak tracking (Elwood, 2020) .
GIS has a tools and techniques is used to create, manage, analyze, and map all types of data. GIS connects data to a map, integrating location data (where things are) with all types of descriptive information (what things are like there). This provides a foundation for mapping and analysis that is used in science and almost every industry. GIS helps users understand patterns, relationships, and geographic context. The benefits include improved communication, efficiency, management, and decision-making. (ESRI, 2024)
GIS systems typically consist of both hardware (computers and peripherals) and software (applications for data manipulation, mapping, and analysis). They facilitate the creation of interactive maps and spatial models, which are essential for understanding complex geographic phenomena and making informed decisions (Ghosh, et al., 2021).

2.1 GIS Techniques

Urban indicators that influence the spread of epidemic infectious diseases include demographic, residential, environmental, and weather factors, as identified in previous studies and in alignment with the guidelines of the UN-Habitat Sustainable Development Goals (SDGs), particularly SDG 3 (Good Health and Well-being) and SDG 11 (Sustainable Cities and Communities) (United Nations, 2016). SDGs 3 and 11 focus on ensuring healthy lives and promoting well-being for all, as well as making cities and human settlements inclusive, safe, resilient, and sustainable.
This research uses GIS techniques as a part of the proposed framework, with the goal of using analytical methods to identify the most vulnerable areas to the spread of infectious diseases. The results will be displayed in the form of maps, and subsequently, the GIS results will be integrated with deep learning method's results.
The GIS works will begin by identifying the key factors that influence the spread of epidemic diseases, including demographic, residential, and environmental factors. These factors will be categorized into main and sub categories, and their relationships with epidemic spread will be analyzed in detail using mathematical methods and the Analytic Hierarchy Process (AHP), relative weights will be assigned to each factor based on its importance and impact on disease transmission.

2.2 GIS Components

Vol.13 Issue 1, Pg.: 1-12
January 2025

A working GIS integrates five key components: hardware, software, data, people, and methods. (Information Resources Management Association, 2016)

Hardware: The physical devices required for GIS operations, including computers, storage devices, GPS units, and printers. Hardware provides the infrastructure needed to collect, process, and display spatial data (Long & et al., 2020).

Software: The programs that process and analyze spatial data. Common GIS software includes ArcGIS, QGIS, and MapInfo. GIS software enables users to visualize data, perform spatial analysis, and manage large datasets. It includes tools for mapping, data query, and advanced analysis (Zhang, et al., 2021).

Key software components:

- Tools for the input and manipulation of geographic information
- A database management system (DBMS)
- Tools that support geographic query, analysis, and visualization
- A graphical user interface (GUI) for easy access to tools

Data: is the geographic (spatial) and non-geographic (attribute) data that is used for analysis. This can include geographic coordinates, environmental factors, population data, and more. Data is the cornerstone of GIS. It can be in vector format (points, lines, polygons) or raster format gridded data like satellite images (Long & et al., 2020).

People: Individuals or teams who design, manage, and analyze GIS systems. People interpret data, make decisions based on spatial analysis, and apply GIS technology in various fields like urban planning, agriculture, and disaster management.

Methods: The processes and techniques used for data collection, analysis, and interpretation within a GIS framework. Methods include spatial analysis (e.g., buffering, overlay), geostatistical modeling, and network analysis to derive meaningful insights from data (Zhang, et al., 2021).

2.2.1 GIS Database

Geo-database: is a database where both spatial and attribute data related to the selected sub-factors are stored and managed (Information Resources Management Association, 2016).

Spatial database management system: The spatial database management system (SDBMA) is an extension of the conventional database management system (DBMS). It is used specially to manage spatial data (Zhang, Y., Liu, & H., 2021).

Structure Query Language (SQL): SQL is used to communicate with the Database and build a simple and complex query, build database stored procedures and triggers (Long & et al., 2020).

2.3 Visualization Techniques and GIS Interface Components

As the aim of integration is to facilitate the understating of result so this part is essential as a communication language between user and the system. GIS visualization function is used to present the spatial results of searched object. A map is usually the result of any geographical project so all needed tools for producing the map should be included in the framework. Then user interface can be developed by any programming language to provide an easy way for the interaction between user and these components (ESRI, 2024).

3. Deep Learning Concept

Deep learning is a subset of machine learning that uses multilayered neural networks, called deep neural networks, to simulate the complex decision-making power of the human brain. Some form of deep learning powers most of the artificial intelligence (AI) applications in our lives today.

The main difference between deep learning and machine learning is the structure of the underlying neural network architecture. “Nondeep,” traditional machine learning models use simple neural networks with one or two computational layers. Deep learning models use three or more layers but typically hundreds or thousands of layers to train the models (Holdsworth & et al., 2024).

3.1 Deep Learning

The researcher will utilize a deep learning as part of the proposed framework. The goal of using deep learning is to use deep learning methods and techniques to identify areas most vulnerable to the spread of epidemic

infectious diseases. Afterward, the results of applying deep learning techniques will be integrated with f the Geographic Information Systems (GIS) results to provide a comprehensive analysis.
The deep learning begins by collecting key weather-related data that influences the spread of epidemic diseases, based on insights from previous studies and expert experiments. The dataset will then be divided into a training dataset, which will be used to train the deep neural network, and a testing dataset, which will be used to evaluate the accuracy and effectiveness of the model.

3.1.1 Deep Neural Network

Traditional neural network architectures, such as the perceptron and Hopfield Neural Networks (HNN), were designed with a simple structure consisting of an input layer and an output layer. However, when the network includes more than two layers (input and output), it is referred to as a deep neural network (DNN). In these deep architectures, each layer processes data based on the output of the previous layer. As the data passes through successive layers, the network is able to extract increasingly complex features and hierarchical representations. This process, known as feature hierarchy, allows the network to learn more sophisticated patterns from the data, as deeper layers can recombine and refine features learned from earlier layers. Consequently, deep networks are better equipped to handle intricate, high-dimensional tasks compared to simpler architectures (LeCun & et al., Deep learning, 2015). These nets allow complex nonlinear relationships see figure 1. (Sangaiah, 2019)

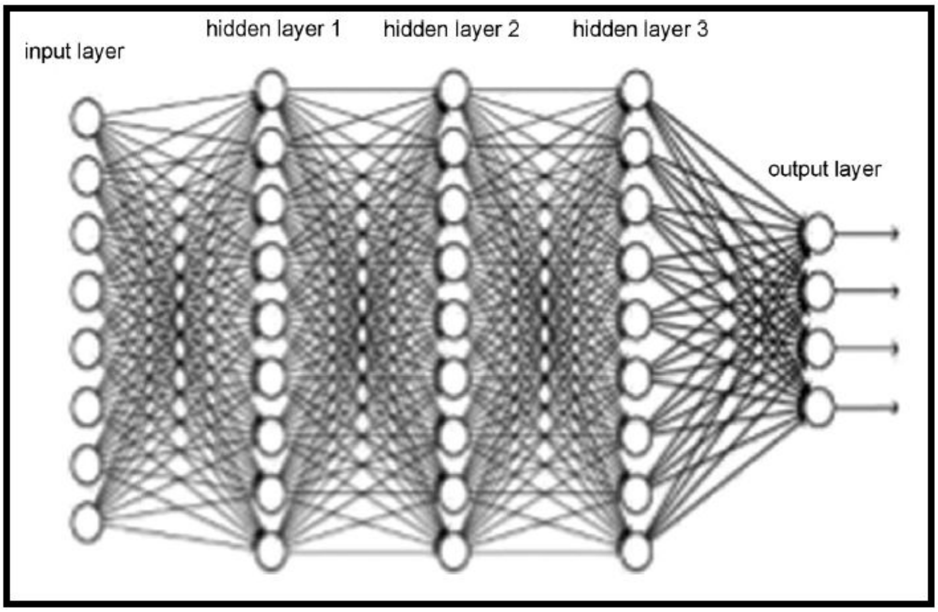

Figure 1, Sample Deep Neural Network, (Sangaiah, 2019)

3.1.1.1 Deep Neural Network Components

Deep Neural networks contains three components as the following (AWS, 2024)
Input layer: An artificial neural network has several nodes that input data into it. These nodes make up the input layer of the system.
Hidden layer: The input layer processes and passes the data to layers further in the neural network. These hidden layers process information at different levels, adapting their behavior as they receive new information. Deep learning networks have hundreds of hidden layers that they can use to analyze a problem from several different angles.
Output layer: The output layer consists of the nodes that output the data. Deep learning models that output "yes" or "no" answers have only two nodes in the output layer. On the other hand, those that output a wider range of answers have more nodes.

Vol.13 Issue 1, Pg.: 1-12
January 2025

4. Main Criteria of the Combined GIS and Deep Learning Techniques

The main criteria of combining GIS (Geographic Information Systems) and deep learning techniques include the following:

1. Data Integration:
GIS excels at spatial data management, while deep learning techniques can analyze complex, high-dimensional data. By integrating both, large-scale and multi-layered geographic data can be processed to gain deeper insights, improving spatial analysis tasks such as land use classification, epidemic spread prediction, and environmental monitoring (Yang, et al., 2021)

2. Feature Extraction and Pattern Recognition:
Deep learning models, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are adept at identifying complex patterns and hierarchies in data. When combined with GIS, they enable the extraction of meaningful features from spatial data, improving the understanding of spatial relationships and trends (Zhu, et al., 2020).

3. Predictive Analysis:
Combining GIS and deep learning allows for advanced predictive analysis, where GIS provides the spatial framework and deep learning models predict future events or patterns, such as the spread of infectious diseases, climate changes, or urban development (Hao, et al., 2021)

4. Scalability and Flexibility:
Deep learning models excel at processing large datasets, and GIS facilitates the management of spatially extensive data. Combining both methods makes it possible to scale the analysis and apply it to various geographic contexts or study areas (Zhang et al., 2021).

5. Dynamic Data Handling:
GIS is designed to manage both static and dynamic spatial data, while deep learning algorithms can handle time-series or sequential data. This combination allows for analyzing temporal changes in spatial data, making it particularly useful for applications like predicting the evolution of disease outbreaks or urban sprawl (Deng et al., 2021).

6. Accuracy and Validation:
GIS provides accurate geographic references, supporting validation and refinement of predictions. By integrating deep learning, the model's accuracy improves as it learns patterns and makes predictions that are validated using GIS-based geographic data (Sun et al., 2020). The combination also ensures that the predicted outcomes have a spatial context, making them more actionable for decision-makers.

7. Optimization:
Deep learning techniques such as reinforcement learning can be used in combination with GIS to optimize decision-making processes. For example, they can help optimize resource allocation, infrastructure planning, and disaster response by using spatially enriched datasets to make informed decisions (Liu et al., 2021).
By leveraging these criteria, the combination of GIS and deep learning enhances the ability to analyze, predict, and optimize spatial processes and systems across various domains, such as epidemiology, urban planning, and environmental monitoring.

5. Building A Proposed Deep Learning Framework Based On GIS

This research proposes a framework for predicting the spatial distribution of epidemic infectious diseases by combining Geographic Information Systems (GIS) and deep learning models. The GIS component utilizes spatial analysis techniques to study and analyze key demographic, residential, and environmental factors that affect the spread of infectious diseases, based on previous studies and expert knowledge. The result is a vulnerability map that ranks regions by their degree of vulnerability, from low to high.
On the other hand, the deep learning classifies vulnerable areas based on dynamic factors like weather data (e.g., temperature, humidity) that influence disease spread, extracted from previous studies and expert

experiments. Training dataset is used to train the deep neural network, while a testing dataset is used to evaluate its performance.

Finally, GIS techniques such as overlay, merge, intersect, and union are applied to integrate the results from both the GIS and deep learning, creating a comprehensive vulnerability map that predicts the spread of epidemic infectious diseases.

The proposed framework shown in **Figure 2**, consists of:

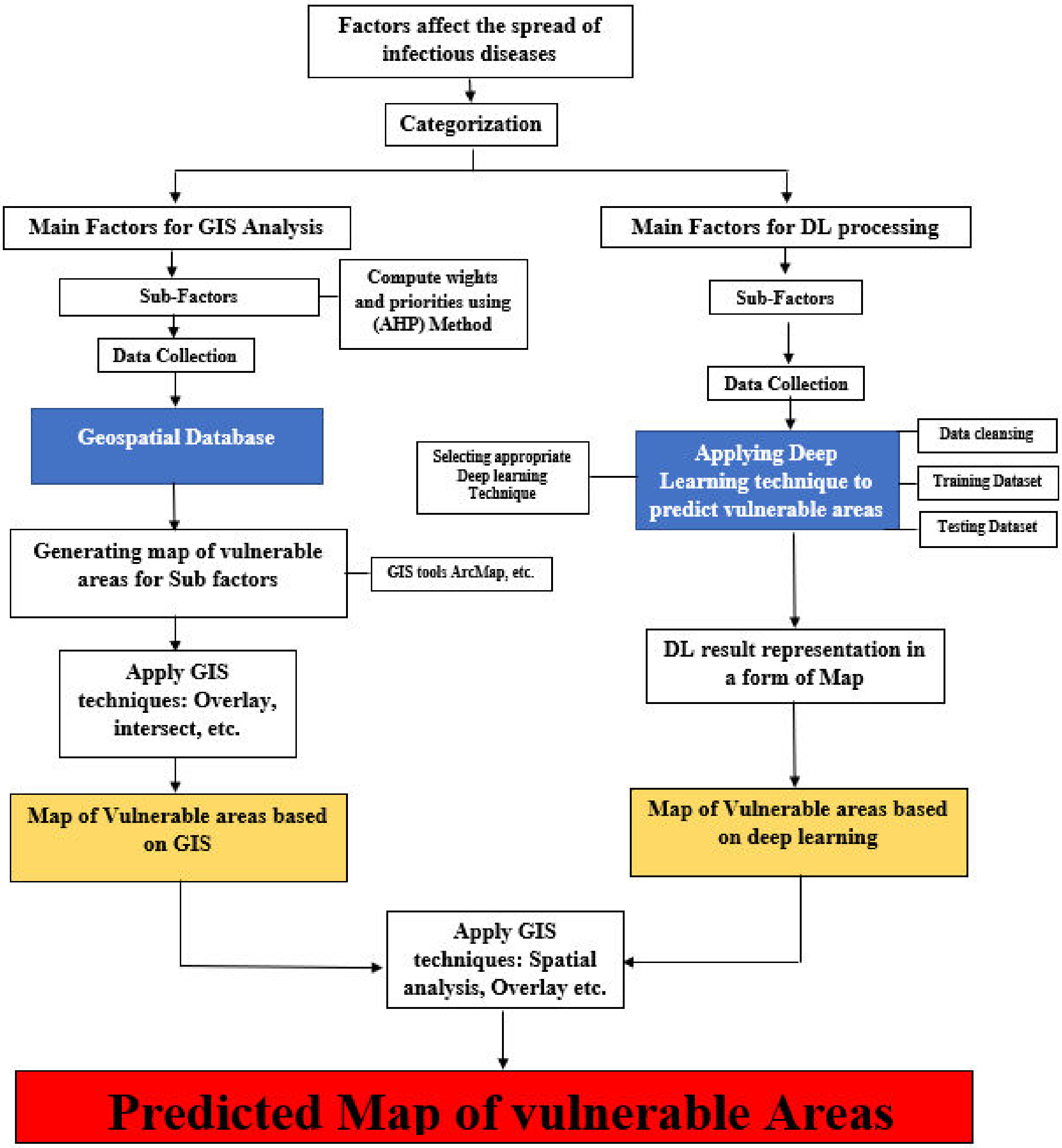

Figure 2, A Proposed Deep Learning Framework Based on GIS to Predict the Spatial Distribution of Epidemic Infectious Diseases.

6. Mechanism of the proposed framework

This section briefly explains the integration between Geographic Information Systems (GIS) and Deep Learning in the proposed framework. In this framework, GIS will be responsible for collecting and analyzing

Vol.13 Issue 1, Pg.: 1-12
January 2025

critical factors, such as demographic, residential, and environmental factors. These factors will be categorized into sub-factors that are relevant for predicting the spread of infectious diseases in urban areas, based on insights from previous studies and expert knowledge. Using this data and create geodatabase, GIS will then generate a vulnerability map for the study areas, highlighting regions at higher risk of disease outbreaks.
While the deep learning will be responsible for learning from large dataset related to weather critical factors such as temperature and humidity, etc. that are relevant for the spread of infectious diseases based on the insights from previous studies and expert knowledge. Using the deep learning techniques, like feed forward deep neural network be trained from the collected dataset and extract patterns related to vulnerability for the study areas which help to predict the spatial distribution of the epidemic infectious diseases. Then put the output in a form of map to apply the GIS techniques like overlay and intersect with vulnerability map generated by GIS to get the final predicted map for the spatial distribution of the infectious diseases in specific period of time for the study areas.

6.1 GIS Mechanism

This part of the study explains how the GIS mechanism of the proposed framework operates, addressing the following key points:
1. Collecting, studying and analyzing factors that are relevant to the spread of the epidemic infectious diseases like demographic, residential and environmental factors etc.
2. Categorizing the selected relevant factors into sub-factors
3. Applying the Analytical Hierarchy Process (AHP) method to compute weights and priorities for the selected relevant sub factors.
4. Data collection for the categorized sub factors.
5. Building the GIS geodatabase for the study areas.
6. Generate vulnerable areas maps for the selected relevant sub indicators.
7. Apply GIS techniques (overlay, intersect) by GIS tools, to generate map of vulnerable areas based on GIS.

6.1.2 Analytical Hierarchy Process (AHP) method used in the framework

The Analytical Hierarchy Process (AHP) is a widely used decision-making method in scientific research, particularly in situations where multiple criteria need to be considered. AHP decomposes a complex problem into a hierarchy of simpler sub-problems, allowing for systematic pairwise comparisons of criteria and alternatives. Decision-makers assign relative weights to each criterion, and these are aggregated to derive a final ranking of options. This method is particularly valuable in multi-criteria decision analysis (MCDA) and has been successfully applied in various fields, including environmental management, healthcare, and urban planning (Saaty & T. L., 2008) (Ho & et al., 2010). Recent studies have extended AHP's applications to areas like sustainable development (Sharma & et al., 2020) and disaster risk assessment (Zhang & al., Disaster risk assessment using AHP and GIS in flood-prone areas, 2021), demonstrating its versatility and robustness in handling both qualitative and quantitative data.
In the proposed framework the researched used the APH method to Analytical Hierarchy Process (AHP) to compute weights and priorities for the selected relevant sub factors.

6.1.2.1 Analytical Hierarchy Process (AHP) structure

The Analytical Hierarchy Process (AHP) is a structured decision-making method used to solve complex problems by breaking them down into simpler components. The basic structure of AHP consists of the following key elements (Saaty & T. L., 2008)
1. Goal (Objective)
The top level of the hierarchy, representing the main objective or decision that needs to be made.
2. Criteria (Factors)
The second level, which includes the factors or criteria that influence the decision. These criteria are used to evaluate the alternatives.
3. Sub-criteria (Optional)
If needed, criteria can be further broken down into sub-criteria, representing more detailed aspects of each criterion.
4. Alternatives

Vol.13 Issue 1, Pg.: 1-12
January 2025

The options or alternatives that are being evaluated, which are located at the bottom level of the hierarchy.

5. Pairwise Comparison

Decision-makers perform pairwise comparisons of the criteria and alternatives at each level, assigning relative importance or preference values using a scale (typically 1-9).

6. Weight Calculation

The pairwise comparisons are used to calculate weights or priorities for each criterion and alternative, typically using eigenvalue methods.

7. Synthesis

The final step involves synthesizing the results by aggregating the weighted scores of the alternatives based on the criteria, producing a final ranking of the alternatives.

6.2 Deep Learning Mechanism

This part of the study explains how the deep learning mechanism of the proposed framework operates, addressing the following key points:

1. Studying changeable factors relevant to the spread of the epidemic infectious diseases like weather factors based on previous studies and experts.
2. Categorize the selected relevant factors into sub factors
3. data collection
4. data cleansing and removing nose from data.
5. Dividing the dataset into training and testing data.
6. Select the appropriate Deep learning Technique based on the case study
7. Applying the Deep Learning techniques to predict vulnerable areas.
8. Add the output data in the form of map
9. generate map of vulnerable areas based on deep learning.

6.3 Integration Between Deep Learning techniques and GIS techniques

This part of the study explains the integration between the vulnerability map resulting from Deep learning and vulnerability map resulting from GIS

1. Applying GIS techniques: Overlay, intersect etc. on the map of vulnerable areas resulting from GIS and the map of vulnerable areas resulting from deep learning.
2. Generate the final predicted map of vulnerable areas.

The proposed framework is highly flexible and can be adapted to various contexts, allowing it to address different epidemic scenarios. While the current implementation focuses on epidemic infectious diseases, the framework can be extended to model other types of epidemics or even diseases with similar transmission patterns. This flexibility stems from the ability to modify the selected factors in the model, which can be tailored to the specific nature of the disease or the geographical and socio-economic characteristics of the study area.

In this framework, GIS techniques are used to analyze and process spatial data, which includes demographic factors (such as population density and age distribution), residential factors (like housing quality and urbanization patterns), and environmental factors (such as air quality, vegetation, and water sources). These factors are categorized into main categories and sub-factors to allow deeper insights into the influences on disease spread. By integrating data on weather conditions, such as temperature, humidity, and precipitation, which are known to impact the spread of certain diseases, the deep learning model can provide more accurate predictions.

Regarding the deep learning techniques, the framework can be easily adapted to incorporate different types of data inputs depending on the use case. For instance:

1. Regression Analysis: When the goal is to predict the magnitude of an epidemic (such as the total number of infections or the peak size of an outbreak), regression models, such as Multi-Layer Perceptron (MLPs), can be employed. These models are designed to predict continuous variables and can provide precise estimates based on input features like temperature, population density, and healthcare infrastructure.

2. Classification Models:

Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) are foundational deep learning models commonly used for classification tasks involving structured data. In cases where the goal is to identify specific high-risk areas (such as predicting whether a given area will experience a high incidence of disease cases), classification models like Convolutional Neural Networks (CNNs) can be used. These models excel at analyzing spatial data, such as satellite images or geospatial grids, and can identify patterns that correlate with epidemic risk. The CNNs are particularly useful in areas where spatial relationships in the data are important for accurate predictions, such as determining the impact of urban sprawl or proximity to water sources.

3. Recurrent Neural Networks (RNNs): for Time-Series Predictions: For time-dependent epidemics, such as predicting the spread of a disease over several weeks or months, Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) networks, can be utilized. These models are designed to process sequential data and capture temporal dependencies, making them ideal for time-series forecasting. RNNs are particularly effective in understanding trends over time, such as seasonal variation in disease transmission or the effect of intervention measures on disease spread.

4. Hybrid Models: The framework also allows the combination of various deep learning models for a more holistic approach. For instance, the output of CNNs can be used as input to an RNN for a two-stage model that first identifies spatial vulnerabilities and then projects the temporal evolution of the epidemic. This hybrid approach leverages the strengths of both spatial and temporal data to provide a comprehensive prediction of the epidemic's spread.

deep learning is capable of handling diverse datasets, including structured data from GIS sources, as well as unstructured data, such as satellite imagery, climate data, and social media signals, which could provide early warnings of disease outbreaks.

Moreover, the framework allows for continuous model retraining. As new data becomes available (e.g., updated population census, real-time weather data, or new outbreak reports), the model can be retrained to improve its accuracy and adapt to the evolving nature of the epidemic. This feature ensures that the framework remains relevant and effective as new information becomes available during an ongoing epidemic.

Lastly, the integration of GIS and deep learning techniques facilitates the creation of vulnerability maps that can be used by decision-makers. By applying spatial and temporal analysis, the framework can produce dynamic, high-resolution maps that predict areas most at risk, enabling authorities to allocate resources and plan interventions effectively."

6.4 Selection of the Deep learning technique for the proposed framework

In the proposed framework, the researcher selects the classification deep learning method, specifically Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP), because these techniques are well-suited for classifying areas most at risk of the spread of epidemic infectious diseases

Feedforward Neural Networks consist of an input layer, one or more hidden layers, and an output layer. The nodes in each layer are fully connected to the nodes in the subsequent layer, and each connection has an associated weight that adjusts during training to minimize the error between the predicted output and the actual target. MLPs are characterized by their ability to model complex nonlinear relationships, making them effective for problems that involve classification, regression, and even function approximation. Typically, they employ activation functions like ReLU or sigmoid to introduce non-linearity into the network, allowing it to learn intricate patterns in data. MLPs are especially useful in tasks where the relationships between inputs and outputs are not straightforward, and they can be trained using backpropagation and optimization techniques such as stochastic gradient descent (SGD). One of the key strengths of MLPs is their flexibility to be applied across various domains, from image and speech recognition to finance and medical applications.

6.4.1 Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) Algorithm

Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) consists of layers of interconnected nodes, or neurons as shown in figure 3, which work together to map input data to output predictions. The structure of an MLP typically includes three types of layers (LeCun, et al., 2015):

1. Input Layer: This layer receives the input data (features) and sends them to the hidden layers. It doesn't perform any computation; it just acts as a conduit for the data.

2. Hidden Layers: These are intermediate layers between the input and output layers, where the actual computation occurs. Each neuron in a hidden layer applies a weighted sum of its inputs, followed by a non-linear activation function (such as ReLU, Sigmoid, or Tanh), and passes the result to the next layer. An MLP can have multiple hidden layers, enabling it to capture more complex patterns and relationships in the data.

3. Output Layer: This layer produces the final result or prediction. For classification tasks, the output is typically a probability distribution over different classes (using a softmax function for multi-class problems or sigmoid for binary classification). For regression tasks, it produces a continuous value.

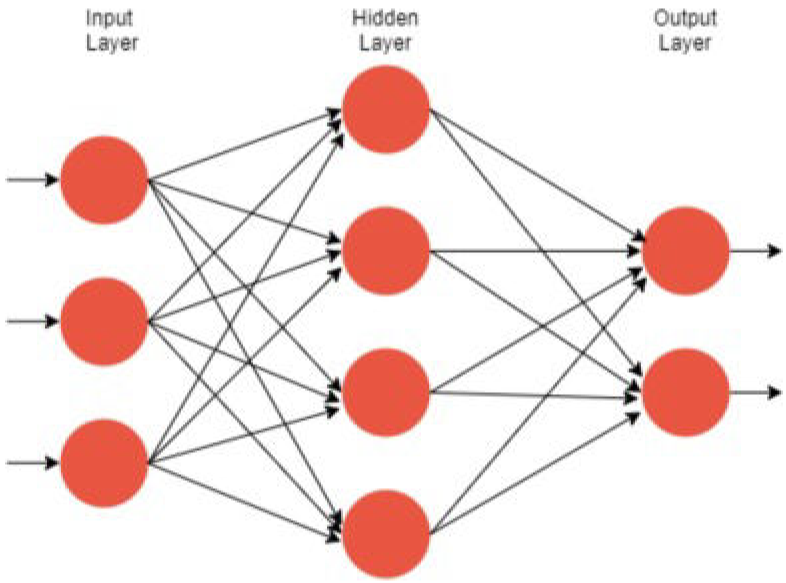

Figure 3, Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) Algorithm, By Researcher

6.4.2 Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) Algorithm Characteristics

The following are the main key Characteristics for the Feedforward Neural Networks (FNN) / Multi-Layer Perceptron (MLP) Algorithm (Chollet & F., 2017):
- **Forward Propagation:** The data moves in one direction, from the input layer to the output layer, through the hidden layers. Each neuron computes a weighted sum of inputs, applies an activation function, and passes the result forward.
- **Backpropagation:** After the forward pass, the model's predictions are compared with the true labels to calculate the loss or error. The backpropagation algorithm is then used to update the weights of the network by minimizing the loss function (usually using gradient descent or other optimization algorithms).
- **Training the MLP:** The network is trained by adjusting the weights based on the error gradient (using the backpropagation algorithm) in order to minimize a defined loss function, which could be cross-entropy loss for classification tasks or mean squared error for regression tasks.
- **Optimization Algorithm:** The most common algorithm used for training MLPs is Stochastic Gradient Descent (SGD) or its variants (e.g., Adam, RMSprop). These algorithms iteratively adjust the weights to minimize the error.

The Feedforward Neural Network (FNN) or Multi-Layer Perceptron (MLP) algorithm is a powerful tool for various classification and regression tasks, capable of learning complex patterns from data. By utilizing multiple layers, MLPs can model intricate relationships in the data, making them highly effective for applications like epidemic prediction, where multiple factors must be considered simultaneously.

7. Conclusion

In this research, a proposed Deep Learning Framework based on GIS was introduced to predict the spatial distribution of epidemic infectious diseases. The framework was designed with flexibility, allowing for effective application in epidemic prediction. Identifying and categorizing the factors influencing epidemic spread is crucial for the framework's effectiveness. The framework’s predictive accuracy depends on the quality of these factors and the data collected. The more accurate and comprehensive the data, the more reliable the framework’s results.

By integrating GIS and deep learning technologies, the framework utilizes deep learning techniques to process large and diverse datasets, identifying patterns and relationships within the data to predict vulnerable areas. alongside GIS techniques that analyze factors related to demographic, residential, and environmental factors relevant to the spread of epidemic infectious diseases to predict the spatial distribution of the diseases. The results of the GIS and deep learning were then integrated to generate actionable insights, assisting decision-makers in implementing strategies to control epidemic spread within the study area.

The proposed framework is designed with flexibility in mind, allowing it to be easily adapted to different contexts. While it has been applied to predicting the spatial distribution of epidemic infectious diseases, the framework is not limited to this specific application. It can be extended to address other types of epidemics, making it highly adaptable for diverse use cases. Additionally, the selected factors influencing epidemic spread can be tailored to suit the unique characteristics of the case study at hand.

The framework also supports the integration of various types of environmental data, such as air quality or land-use patterns, providing further flexibility. Furthermore, it accommodates multiple deep learning techniques depending on the specific needs of the problem. For example, regression methods can be employed for predicting the size of an epidemic, while classification techniques can be used for identifying high-risk areas. Depending on the nature of the data, deep learning methods such as Convolutional Neural Networks (CNNs) for spatial analysis or Recurrent Neural Networks (RNNs) for time-series forecasting can be seamlessly incorporated into the framework for more accurate predictions.

References

Liu, Z., Shi, P., & Zhang, L. (2021). Optimal resource allocation in disaster management using GIS and deep learning. Computers, *Environment and Urban Systems*, 85, 101559. https://doi.org/10.1016/j.compenvurbsys.2020.101559.

AWS. (2024). What is deep learning. Retrieved from AWS: https://aws.amazon.com/what-is/deep-learning/

Chollet, & F. (2017). Deep Learning with Python. *Manning Publications*.

Chung, A., Choi, W. &., J., Kim, & Y. (2020). Deep learning-based prediction of infectious disease outbreaks. Nature Communications, 11(1), 1043 . https://doi.org/10.1038/s41467-020-14976-1.

Deng, J., Zhang, S., & Liu, M. (2021). Integrating deep learning and GIS for time-series spatial analysis of environmental changes. *Environmental Monitoring and Assessment*, 193(12), 1-12. https://doi.org/10.1007/s10661-021-09025-5.

Elwood, e. a. (2020). The role of geographic information systems in spatial data analysis and decision-making. *Annals of the American Association of Geographers*, 110(1), 8-27. https://doi.org/10.1080/24694452.2019.1671952.

ESRI. (2024). GIS Overview. *Retrieved from ESRI*: https://www.esri.com/en-us/what-is-gis/overview

Ghosh, e. a. (2021). Geographic Information Systems (GIS) and their applications in spatial data analysis and decision-making. *International Journal of Geographical Information Science*, 35(4), 746-765. https://doi.org/10.1080/13658816.2020.

Goodfellow, I., Bengio, Y., Courville, & A. (2016). Deep Learning. MIT Press.

Hao, W., Wang, X., Guo, & X. (2021). Integration of GIS and deep learning for predictive analysis of epidemic diseases: A case study on COVID-19. *Journal of Geographic Information Systems*, 11(1), 1-13. https://doi.org/10.1080/13658816.2021.1877435.

Ho, & et al. (2010). A review of decision-making models and approaches for supporting product design and development. *Computers in Industry*, 61(3), 173-181.

Holdsworth, & et al. (2024, Jun 17). What is deep learning ? *Retrieved from IBM*: https://www.ibm.com/topics/deep-learning

Information Resources Management Association. (2016). Geospatial Research : Concept, Methodolgies and Applications. USA: *Information Science Refrence* (an Imprint of IGI Global), p.799.

LeCun, & et al. (2015). *Deep learning. Nature*, 521(7553), 436-444. https://doi.org/10.1038/nature14539.

LeCun, Y., Bengio, Y., Hinton, & G. (2015). *Deep learning. Nature*, 521(7553), 436-444.

Long, & et al. (2020). GIS technology and its applications in urban planning. *A comprehensive review. Sustainability*, 12(13), 4195. https://doi.org/10.3390/su12134195.

Mark, M. D. (2000). Geographic Information science: Critical issues is an emerging across-disciplinary research domain. *Journal of the Urban and Regional Information Systems Association*, 45-54.

National Geographic. (2024). GIS (Geographic Information System). *Retrieved from national geographic*: https://education.nationalgeographic.org/resource/geographic-information-system-gis/

Ruktanonchai, N. W., Rojas, D. P., &Tatem, & A. J. . (2020). GIS and machine learning for epidemic forecasting: Predicting disease outbreaks in real-time. *Computers, Environment and Urban Systems*, 79, 101405. https://doi.org/10.1016/j.compenvurbsys.2019.1014.

Saaty, & T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), 83-98.

Sangaiah, A. K. (2019). Deep learning and Parallel Computing Environment for Bioengineering Systems. *Mara Conner*.

Sharma, & et al. (2020). Sustainability, 12(9), 2856.

Sharma, Ram, Prasad, & Binda. (2006). Mathematical Modelling In Geographical Information Systems Global Positioning System and Digital Cartography. New Delhi: Ashok Kumar Mittal, *Concept Publishing Company*, p. 195 - 197.

Sun, M., Li, Y., & Zhang, Y. . (2020). Enhancing spatial prediction of epidemics using GIS and deep learning: A systematic review. *Journal of Epidemiology and Global Health*, 10(3), 219-227. https://doi.org/10.1016/j.jegh.2020.06.001.

United Nations. (2016). SDG Indicators: Metadata Repository. *United Nations Statistics Division*. Available at: https://unstats.un.org/sdgs/metadata/ https://unstats.un.org/sdgs/metadata/.

Yang, X., Zhang, Y., Chen, & X. (2021). GIS and deep learning for urban development and environmental monitoring. *Journal of Geographical Information Science*, 35(6), 1125-1138. https://doi.org/10.1080/12345678.2021.1767639.

Zhang, & al., e. (2021). Disaster risk assessment using AHP and GIS in flood-prone areas. *Natural Hazards*, 105(1), 39-59.

Zhang, e. a. (2021). Geographic Information Systems (GIS) and their role in spatial data analysis and decision-making. *Journal of Spatial Information Science*, 2021(1), 1-10. https://doi.org/10.1007/s41651-021-00088-5.

Zhang, Y., Liu, & H. (2021). Geographic information system (GIS) and its applications in environmental management. *Environmental Science and Pollution Research,* 28(7), 8105-8117. https://doi.org/10.1007/s11356-020-10799-3.

Zhou, M., Xie, Z., &Li, & Z. (2021). Integrating deep learning and GIS for epidemic prediction and management. *International Journal of Environmental Research and Public Health*, 18(6), 2823. https://doi.org/10.3390/ijerph18062823.

Zhu, X., Yang, & X. (2020). Combining deep learning with GIS for intelligent urban and environmental monitoring. *Remote Sensing*, 12(6), 946. https://doi.org/10.3390/rs12060946.