Early Warning Panel for Epidemiological Outbreaks

Diego M. Olmedo Giménez[1], Nonia I. Acosta Britos[1], Juan V. Bogado Machuca[2]

[1]*National University of Caaguazú, Faculty of Sciences and Technology*
Coronel Oviedo/Caaguazú, Paraguay
diolmedo@fctunca.edu.py, noacosta@fctunca.edu.py
[2]*National University of Caaguazú*
Coronel Oviedo/Caaguazú, Paraguay
jbogado@unca.edu.py

Abstract. Dengue has become a serious health problem in Paraguay, as it occurs in tropical and subtropical regions. With the help of accurate forecasts, preventive measures can be taken to combat this disease. In response to this problem, the present research has made it possible to solve the above, developing a web page for Early Warning of Epidemiological Outbreaks in Paraguay, whose sample used for this research is based on Dengue. It presents graphs of epidemiological indicators such as incidence, temperature variables and predictions, so that decision makers can analyze the trends of Dengue in Paraguay in its 17 departments and districts. The data collected for the analysis correspond to the years 2009 to 2013. The developed dashboard offers an intuitive interface that displays a map of dengue cases, along with previously selected relevant indicators. In addition, an artificial intelligence model was implemented to make predictions of dengue cases. The Panel was evaluated in three main sections: usability, relevance and performance, through a survey of a group of experts in epidemiology and infectious disease control. The results of this evaluation revealed a good acceptance by the experts.

Keywords: Early warning, Artificial Intelligence, predictive Model, epidemiological indicators, disease Surveillance.

1 Introduction

Paraguay is considered an endemic country for certain arboviruses from the epidemiological point of view. Dengue, a disease that has become a major health problem in the country, appears in tropical and subtropical regions, with records of cases since 2009. Likewise, there are records of Chikungunya since 2013 and Zika since 2015 [1].

In this paper, we present information on dengue surveillance through a dashboard that displays epidemiological indicators in interactive graphics. This will allow decision-makers to visualize information in real time and interact with it, thus improving disease responsiveness. In addition, predictive graphics are displayed, justifying the use of artificial intelligence.

Artificial intelligence methods are normally used to make predictions, classify targets, cluster and optimize human work and minimize uncertainty for decision making, but despite the advantages they offer, they are not entirely friendly or accessible to most people, and the proposed techniques found are not directly usable by medical professionals or managers who normally have to make decisions involving the health of a country. That is why we use a trained artificial intelligence model, the LSTM model, since it is one of the most widely used to predict time series and can be extended to any endemic disease for which there is sufficient data for deep learning of an artificial intelligence model, which will allow us to draw plans to combat and cope with the disease in our country.

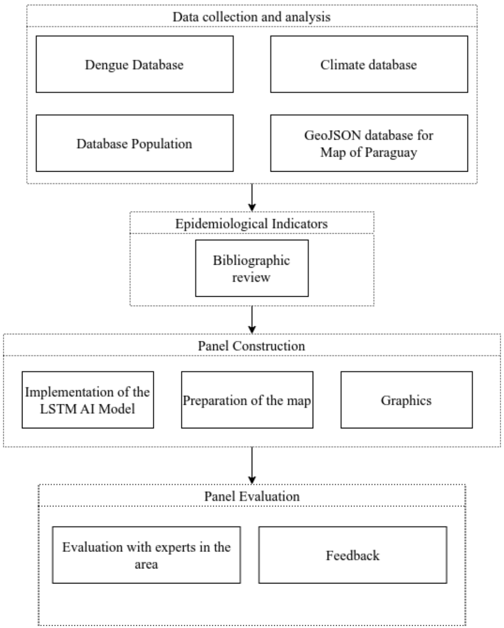
2 Methodology

2.1 Materials and methods for the Creation of an Early Warning Panel

The methodology applied in this work is developed from data collection to panel evaluation. Initially, data on dengue cases were obtained from a specific database. In addition, relevant climate data were collected and the databases were populated with demographic information essential for the analysis. A GeoJSON file was used to create detailed maps of Paraguay, which facilitated the geospatial representation of the data.

With the data collected, epidemiological indicators were identified and used through a comprehensive review of the scientific literature. Subsequently, a trained artificial intelligence model based on Long Short-Term Memory (LSTM) was implemented to make predictions. The elaboration of detailed maps and the generation of visual graphs allowed a better understanding and analysis of the data.

Experts in epidemiology and related areas provided valuable comments to improve and fine-tune the analysis. This comprehensive and systematic methodology ensured a rigorous approach in the development of the project, from data collection to the application and evaluation of the predictive model.



2.2 Data

The data sample used for this research was from the Comidenco project, specifically these databases are related to dengue cases distributed by: year, epidemiological week and temperature, from 147 districts of Paraguay from 2009 to 2013.

It should be noted that, in this project, the cases were organized by districts on a weekly basis, coding the departments and districts. In addition, for the data on the number of inhabitants of each district, data from the population projections of the National Institute of Statistics (INE) [5], from 2000 to 2025, were used. For the interactive map of Paraguay, INE datasets were also used, specifically geographic codes from the 2012 National Population and Housing Census[6].

The nomenclature of the districts is based on that of the INE, i.e., we use the numerical identifier of each district associated with its department, for example: the city of Coronel Oviedo is number 1 in the list of districts of Caaguazú, which is department number 5, therefore the code is as follows: 0501. The details of this database are shown in Table Nº 1.

Table 1. Base structure

Field	Description	Type
year	Year corresponding to epidemiological Week	Numeric
week	Epidemiological week	Numeric
district_id	District identifier	Numeric
district	District name	Character
tmed	Average temperature per week	Float
cases	number of cases	Numeric
department_code	Department identifier	Numeric
department	Department name	Character
population	number of inhabitants per district	Numeric

2.3 Epidemiological Indicators.

Epidemiological indicators provide us with information on the epidemiological situation of Dengue in Paraguay, so that we can observe its course and prepare measures for the control of the disease. The criteria for their selection were the following: articles that mention epidemiological indicators and the way in which they are calculated.

Below are the indicators that we compiled through a bibliographic review and that are found in our Early Warning Panel:

The incidence of Dengue: Data on the number of dengue cases and the population of the selected district are used to calculate the incidence.

$$Icd = 100 \times \frac{c}{p}$$

Temperature-dependent indicators: the following epidemiological indicators use mean temperature variables corresponding to the selected district with its epidemiological week

a) The biting rate:

$$b_t = \begin{cases} 0.000202T(T - 13.35)\, 40.08 - T & (13.35°C \;\; T \;\; 40.08\,°C) \\ 0 & (T < 13.35°C,\ T > 40.08\,°C) \end{cases}$$

b) The probability of infection from human to mosquito per bite.

$$b_h(T) = \begin{cases} 0.001044 \times T(T - 12.286) \times \sqrt{32.461 - T} & (12.286°C \leq T \leq 32.461\,°C) \\ 0 & (T < 12.286°C,\ T > 32.461\,°C) \end{cases}$$

c) The probability of infection from mosquito to human per bite.

$$b_m T = \begin{cases} -0.9037 + 0.0729T & (12.4°C \;\; T < 26.1\,°C) \\ 1 & (26.1°C \;\; T \;\; 32.5\,°C) \\ 0 & (T < 12.4°C,\ T > 32.5\,°C) \end{cases}$$

2.4 Model LSTM.

For the implementation of the prediction in our Early Warning Panel we defined the LSTM model, by its acronym Long Short-Term Memory proposed by Hochreiter and Schmidhuber in 1997. It is one of the most advanced and successful deep learning architectures for time series prediction, handwriting recognition and speech analysis [4] and with these features we found that this model meets the objectives we need for prediction in our panel, so the mentioned model has been implemented.

2.5 Results

In this section we present the way in which the selected indicators were presented, as well as the predictions and the interactive map of cases of Paraguay.

2.5.1 Incidence

Incidence prediction is important for the effective use of limited resources for vector control and contingency measures [7].

According to[8] for the calculation of incidence, once the time series of each city is obtained, the incidence is calculated.

The result of the Incidence by department graph can be seen in Image 1.

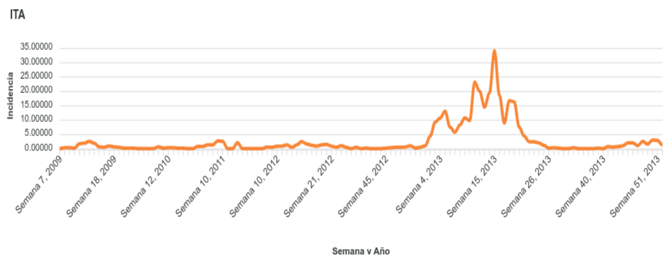
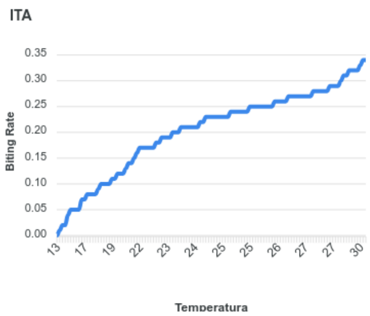


Image Nº 1: Incidence of the department of Ita.

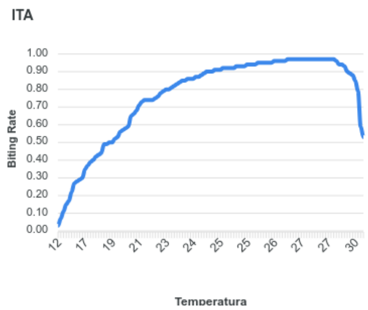
2.5.2 Temperature

It should be noted that climatic factors such as temperature and precipitation significantly affect the life cycle of dengue mosquitoes; therefore, these factors should be included in mathematical models of dengue transmission. It is important to include more realistic parameters for dengue transmission[9]. These include: biting rate, the probability of infection from human to mosquito per bite and the probability of infection from mosquito to human per bite.

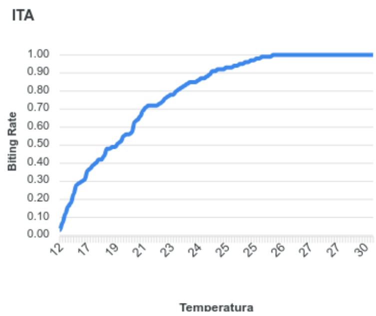
The temperature plots obtained in our panel are shown in Image (a), (b) and (c).



(a)



(b)



(c)

It should be noted that basic statistical information has also been included such as: the number of cases represented on a map of Paraguay, with a filter by department with its corresponding district, also can be entered by date range, selecting the date from and date to,, this interactive map also provides information on the number of cases with a heat map, while the department has more cases, the color is darker, allowing the analysis at a glance of the user, in addition to the geolocation of the entire country with respect to cases of dengue, as can be seen in Image 5.

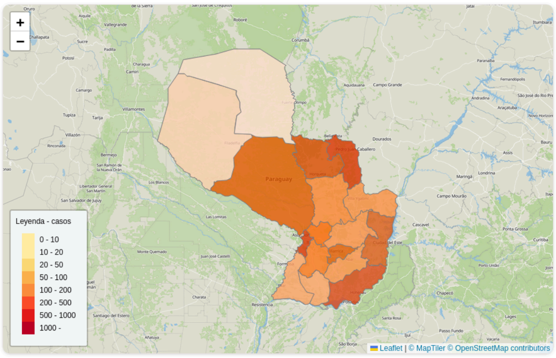


Image Nº 5: Heat Map with Dengue Cases.

Another section of statistical information that we developed in the Panel is the percentage of cases that are subdivided into the 17 departments, indicating the percentage that corresponds to Paraguay, the result can be seen in image 6.

Image Nº 6: Percentage of cases by department.

FastApi is used to obtain the prediction, which is a framework for building APIS Rest Full with Python, since the predictive model used requires Python libraries for its implementation (Image 7).

Image Nº7: Predictions

2.6 Structure of the early warning web panel

You are the sole responsible for making sure that you have the right to publish everything in your paper. If you use material from a copyrighted source, or from other authors, you may need to obtain permission from the copyright holder or the respective authors. An "Authorship statement" must be placed at the bottom of your paper, immediately before the list of references, as indicated further below.

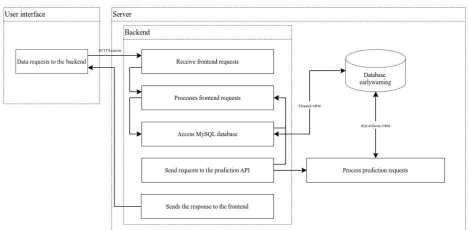


Figure 1. Flowchart showing the structure of the Early Alert Panel

2.7 Algorithmic representation of the use of the Predictive Model using an API

For obtaining predictions, FastAPI is utilized, a modern and high-performance framework for building RESTfull APIs using Python. The choice of FastAPI is due to its capability to efficiently handle asynchronous requests and its compatibility with the necessary Python libraries for implementing the predictive model. This approach allows for smooth integration and robust deployment of the artificial intelligence model in a production environment.

Data retrieval: dengue cases are obtained from the database according to a start date and an end date, in addition to the corresponding district ID.

Data processing: The column of cases is extracted and converted into a DataFrame, which is then stored in a variable.

Model loading: The predictive model is loaded into the environment.

Prediction: The prediction is performed using the loaded model and the cases contained in the DataFrame.

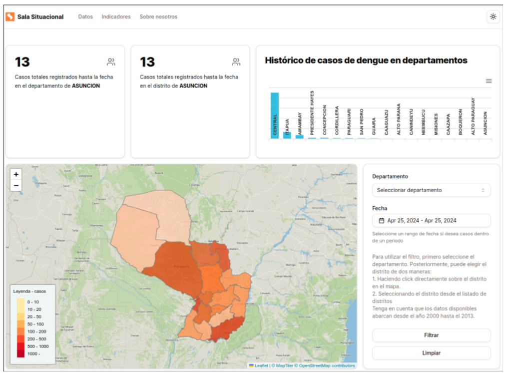
```
Algorithm 1: Algorithm with Pre-trained Model
Input: Date range and corresponding district
Output: Array of predictions
Data: "earlyalert" database of the denguecases table

/* API endpoint to retrieve dengue case predictions for a
   given date range and district                                  */
Def getPredictions(startDate, endDate, district):
    query ← Select cases column from Dengue-Data table
                ((dataDate ≥ startDate)∧
    conditions ← (dataDate ≤ endDate))∧
                idDistrict = district
    Apply conditions to query
    results ← ExecuteQuery(query)
    casedata ← ExtractData(results)
    predictions ← makePredictions(caseData)
    return predictions

/* Helper function to generate predictions using a
   pre-trained model and input case data                          */
Function makePredictions(cases):
    dataInput ← CreateDataFrame(cases, column = 'cases')
    model ← LoadModel('path/of/model')
    predictions ← model.Predecir(dataInput['cases'])
    return predictions.flatten(∅)
```

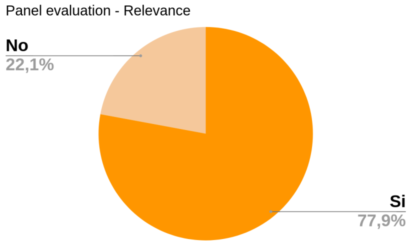
Figure 2. The prediction algorithm

An image of the early warning panel is shown below.

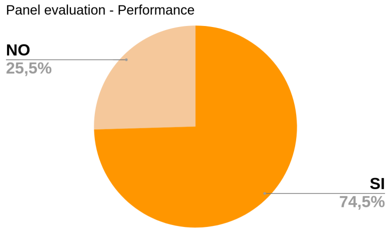


2.8 Evaluation of The Early Warning Panel

For the evaluation of our Panel, a survey of potential users was developed to assess the usability, relevance and performance of the system. Experts in epidemiology and infectious disease control, with experience in dengue case management, were selected. A representative sample was sought that included epidemiologists, tropical disease physicians, data analysts, and public health officials.



77.9% responded that the Panel is relevant and 22.1% responded that it is not.



74.5% responded that the panel performs well in terms of loading graphics, filters and the panel in general, while 25.5% responded that it does not.

3 Conclusions

In conclusion, we emphasize the ongoing importance of Health Surveillance to comprehensively address diseases throughout the country. The Early Warning Panel for epidemiological outbreaks represents a first step towards an interactive analysis based initially on dengue, with potential for extension to other endemic diseases based on the data collected. The identification of relevant epidemiological indicators, the development of an intuitive web panel with predictive models of dengue cases, and the evaluation of its usability, relevance and performance by expert epidemiologists have all been accomplished. This study contributes significantly to the advancement of knowledge in dengue epidemiology and provides a practical tool to improve the management of this disease, which is expected to benefit public health professionals and future research in this crucial field.

Acknowledgements. We extend our deepest gratitude to all the institutions that allowed us access to the necessary data, thus enabling us to achieve our research objectives. We also express our gratitude to the Faculty of Science and Technology and the Rectorate of the National University of Caaguazú for providing us with a solid academic foundation and for their continuous support throughout the conduct of this research.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] «Salva Vidas Sin Criaderos. Dengue, Zika, Chikungunya». Accedido: 12 de septiembre de 2022. [En línea]. Disponible en: https://www.mspbs.gov.py/dengue-zika-chikungunya.html

[2] C. García Pérez y P. Alfonso Aguilar, «Vigilancia epidemiológica en salud», *Rev. Arch. Méd. Camagüey*, vol. 17, n.º 6, pp. 121-128, dic. 2013.

[3] G. Coutin Marie, «Utilización de modelos ARIMA para la vigilancia de enfermedades transmisibles», *Rev. Cuba. Salud Pública*, vol. 33, n.º 2, Art. n.º 2, jun. 2007.

[4] «Estimación adelantada del crecimiento regional mediante redes neuronales LSTM - Dialnet». Accedido: 12 de septiembre de 2022. [En línea]. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=7867482

[5] «INE::Instituto Nacional de Estadística». Accedido: 6 de abril de 2024. [En línea]. Disponible en: https://www.ine.gov.py/microdatos/datos-abiertos-indicador.php?ind=16

[6] «CÓDIGOS GEOGRÁFICOS - CENSO NACIONAL DE POBLACIÓN Y VIVIENDAS 2012 - PARAGUAY | Datos.gov.py». Accedido: 20 de abril de 2024. [En línea]. Disponible en: https://www.datos.gov.py/dataset/c%C3%B3digos-geogr%C3%A1ficos-censo-nacional-de-poblaci%C3%B3n-y-viviendas-2012-paraguay

[7] P. Dayama y S. Kameshwaran, «Predicting the Dengue Incidence in Singapore using Univariate Time Series Models», p. 8.

[8] J. V. Bogado, D. H. Stalder, C. E. Schaerer, y S. Gómez-Guerrero, «Time Series Clustering to Improve Dengue Cases Forecasting with Deep Learning», en *2021 XLVII Latin American Computing Conference (CLEI)*, oct. 2021, pp. 1-10. doi: 10.1109/CLEI53233.2021.9640130.

[9] H. Lee, J. E. Kim, S. Lee, y C. H. Lee, «Potential effects of climate change on dengue transmission dynamics in Korea», *PLOS ONE*, vol. 13, n.º 6, p. e0199205, jun. 2018, doi: 10.1371/journal.pone.0199205.