Prediction of Dengue Cases in Paraguay Using Artificial Neural Networks

V. Ughelli[1], Y. Lisnichuk[1], J. Paciello[1] and J. Pane[1]
[1]Computer Engeneering, Universidad Nacional de Asuncion, San Lorenzo, Central, Paraguay

Abstract— *Dengue Fever is a disease that has grown worldwide in the last years. Several studies show that weather condition is related to the disease, however, as far as we know there is no study in Paraguay that reveals this relation. In this work an analysis of the influence of variables is performed and the efficiency of the neural networks is used to predict the number of disease cases. Thus, this work proposes finding a prediction model of the number of dengue cases with up to 4 weeks of anticipation for districts of Paraguay finding most influential climatic variables. In addition, a variable selection and prediction method that can be used for any geographical region was developed showing promising results.*

Keywords: Dengue Fever, Artificial Neural Networks, Prediction

1. Introduction

The World Health Organization defines Dengue Fever as a viral disease transmitted by mosquitoes that has spread rapidly in recent years in all regions [1]. In Paraguay it has been endemic since 2009 and reached its maximum peak of cases in 2013[3]. This disease can be fatal and has already claimed thousands of lives worldwide. Currently, about 40% of the world's population is at risk to get dengue fever [2].Therefore, their study and analysis become of vital importance. The lack of tools to predict and quantify the spread of the disease is a great difficulty in the adequate estimation of resources and necessary actions by the institutions responsible for ensuring the population's health. At the global level there are several studies that analyze the behavior of the transmitting vector and the disease. Through these analysis predictions are made about the approximate number of cases that could occur in a given time. In Paraguay, however, although there are studies related to Dengue Fever, there is still no tool or model that is able to identify the correlated variables (also called co-variables) that affect and predict the number of cases of the disease. Ojeda et al. [8] proposes a classification model that allows to determine the occurrence of outbreaks of dengue a week in advance. They identify that the co-variables that affect the outbreak are: the population, incidence (calculated as the number of cases per 10.000 individuals of the population for the considered region) and number of cases of the previous week, without identifying other co-variables that could potentially be related.

An outbreak is defined as the occurrence of cases in excess of a disease compared to what is normally expected in a given community, geographical area or time [1]. However, Ojeda et al. [8] mention that the quartiles technique used by the Health Surveillance Office of the Ministry of Public Health and Social Welfare (known in Spanish as DGVS-MSPBS) to classify an outbreak or not in the country in a given week is not effective for all areas of the country. The authors recommend predicting the number of cases of the disease rather than the existence or non-existence of an outbreak.

Weather conditions directly and indirectly affect the development of dengue vectors. For example, the temperature influences vector development rates, mortality, behavior, and controls viral replication of the mosquito.The variability in precipitation influences the habitat development for the vector. Indirectly, rain, temperature and humidity influence the coverage and use of land, which can promote or prevent the growth of vector populations. [14]

This work proposes the implementation of a predictive model of the number of dengue cases and the identification of other co-variables that affect this number. It also proposes to make forecasts with more than a week in advance, thus increasing the time available for the authorities to prepare an action plan, since only one week of time may not be enough.

Artificial Neural Networks are used to implement the proposed model. They are able to find patterns between the nonlinear input variables and the variable to predict. To optimize the values of parameters necessary to construct a Neural Network a tool is used to evaluate different combinations of parameters values, under certain heuristics, in order to obtain adequate values to make predictions of the number of dengue cases. It is also implemented a tool that is capable of: I) Analyze the influence of climatic variables related to dengue cases and, ii) determine the model that obtains a high precision in the prediction of number of cases when used in the Neural Network. With this tool it is possible to easily identify the co-variables related to this disease, independently of the geographic region.

2. Problem Definition

The main problem to be solved is the lack of a predictive model of the number of dengue cases with a reasonable anticipation time that will allow the authorities to make better decisions in a timely manner.

The specific problems to be solved are as follows:

- The existing dengue prediction model for Paraguay, based on the definition of an epidemiological outbreak, has the problem presented in the previous section.
- The current time for decision-making in the analysis of dengue cases is limited to a maximum of one week in advance, which may not be enough.
- The determination of the input co-variables that affect the prediction model is a non-trivial process due to the large number of possible combinations.
- The determination of the optimal parameters values for a neural network destined to make predictions of dengue cases is a complex task due to the large number of possible combinations.
- The lack of a web application that allows the authorities to predict the number of dengue cases with a reasonable time of anticipation and to identify the variables that influence the predictions.

3. Related Works

Within the existing literature, the prediction of dengue cases is a problem that has been tackled from different perspectives and using different techniques according to their region of study.

The World Health Organization considers that one of the most important macrofactors in the increase of the number of dengue cases is climate change, and indicates that global warming and the "El Niño" phenomenon could alter the Ecosystems and generate better opportunities for the survival of pathogens and vectors [11]. This is why several studies focus on the use of climatic variables to predict the number of cases of the disease.

Gultekin et al. [4] proposed the use of artificial neural networks for the prediction of dengue cases in Singapore using the average temperature, the relative humidity and the total amount of rain per week, obtaining good results, without specifying the number of weeks of anticipation of the input variables with respect to their output variable (number of cases). However, this definition is important to know because it determines how much time a variable affects the number of cases to predict.

Nishanthi et al. [5] also used artificial neural networks with average temperature and humidity with lags of 1 and 4 weeks, cumulative rainfall with a lag of 4 weeks and number of cases from the previous week as inputs to predict the number of cases in Sri Lanka. For Ling Hii [6], the predominant factors in predicting cases are the average temperature and accumulated rainfall with lags from 9 to 16 weeks for Singapore. The co-variables and network architectures found by both works serve only for the data of their countries.

In Colombia, Rua-Uribe et al. [12] proposed a statistical model based on time series to predict the incidence of dengue in Medellin, finding that precipitation was the climatic variable statistically more significant associated with incidence of dengue, with an anticipation of 20 weeks, neither humidity nor temperature showed a statistically significant association.

With this, it can be observed that the prediction model and its input variables vary between regions, with no single model of prediction at the global level since each region has its own characteristics and conditions.

In Paraguay there are some related studies, Baez [7] proposes a predictive model of dengue foci using a simulated environment, which allows to perform complex analysis of the spatial reality quickly, generating regionalized information to determine the infestation levels corresponding to the geographical area studied. The work was carried out under a simulated environment and its results were not compared with actual data on dengue cases or larval infestation in Paraguay.

On the other hand, Ojeda et al. [8] proposes a standard model for the publication of data on dengue cases, and introduces a classification model of dengue outbreak with decision trees, using outbreak data, number of cases from the previous week, population and incidence to determine whether or not an outbreak will occur in the next week. Climatic data or other variables that are not dependent on the number of cases are not used in their final model since they included non-favorable experimental results.

4. Proposed Solution

The model proposed in this work aims to find out the variables that influence the propagation of the dengue transmitter mosquito and, consequently, in the number of dengue cases.

4.1 Data Collection

We attempted to use data from different sources and natures correlated to dengue to perform the prediction of cases. The table 1 summarizes the data that were used in the prediction model.

Table 1: Data Collection

Data	Source
River level	The daily level data of rivers of Paraguay were extracted from the site DINAC and then used the maximum levels achieved each week by district from 2009 to 2015.
Climatic variables	Accumulated rainfall in millimeters and average, maximum and minimum daily histories of temperature in degrees Celsius and relative humidity in percentage, were obtained from the Wunderground site.
Dengue cases	Number of cases per district per epidemiological week were obtained from the standard report proposed by Ojeda et al [8], which were obtained from the records of acute febrile syndrome collected by the Health Surveillance Office from 2009 to 2015.

The climatic data were aggregated using the mean value to obtain a value per epidemiological week as data from dengue cases have this level of granularity.

4.2 Feature Selection

Feature selection is the process of selecting a subset of significant variables for using in the prediction model. Feature selection serves two main purposes: First, it can make the training more efficient reducing the variables number and second, feature selection often increases the prediction accuracy by eliminating noise features. A noise feature is one that, when added to the training data, increases the prediction error on new data.[13]

In the prediction of dengue cases, a previous feature selection process is important since initially there is a large possible combination of variables to be used. For each of the input variables, it is necessary to consider the lag it affects the number of cases, so, for example, the maximum temperature could affect to the number of cases with $1, 2, 3...n$ weeks in advance, apart of its combination with the other variables and their possible lags.

The first step in the selection was to eliminate redundant variables to avoid over-information training. The over-information occurs when two or more variables give the same information to the model, so it is convenient to use only one of them. To perform this elimination, the Pearson correlation analysis was used among the possible available variables. [15] The first conclusion in this analysis was that variables related to altitude level of the river were highly correlated with the temperature variables. Also, the temperature variables had higher correlation with the variable to be predicted, the number of cases, so that the variable of altitude level of the river were finally not used.

Then the Pearson correlation analysis was performed for the other variables with respect to each other and with respect to the variable to be predicted with different lags for each of them.

The following conclusions were also reached:

- The maximum lag in which the variables still have a correlation with the number of cases is twelve weeks
- Pearson's correlation was low for most variables, so we can conclude that the correlation between the climatic variables and the number of cases is not linear.
- Climatic variables of close lag between them have high correlation, so it is preferable to use only one of each type (temperature, humidity, rainfall) to avoid overinformation.

Given these conclusions an heuristic was developed that can be reused with data of any country and to implement previous selection of variables process, presented in Algorithm 1 For each possible climatic variables: the average, maximum and minimum temperature, humidity and accumulated rain along with their lags from 1 to a specified n, in our case 12, the two variables of each type are selected with the highest Pearson's correlation with the number of cases of the week to be predicted. Once the two variables are found for each type of climatic variable, we have 6 possible variables and considering the number of cases variable from the previous week, we have 7 possible variables in total. Then, these variables are combined with each other, ensuring that there is no over-information, thus generating at most 27 possible models. The previous analysis does not consider the relationship between the combination of variables and the number cases of dengue, only between each of them individually with the cases. To include this relationship these, models were then tested by the Neuronal Network to find the best final model.

Algorithm 1 Selection of climatic variables to predict number of dengue cases

Require: Weekly climatic data and cases per district
Ensure: Sub-sets of combined variables to use in prediction
for Variable (humidity, temperature, rain) delay 1 ... n **do**
    variables += Find two better variables with higher Pearson correlation with number of cases
end for
combinations = combine(variables,number of cases of the previous week) (number of cases, humidity1, humidity2, rain1, rain2, temperature1, temperature2)
return combinations

4.3 Predictive model Implementation

After identifying the co-variables with the highest correlation respect to the number of dengue cases, the next step is to find a method for predicting the number of cases of dengue based on the values of these co-variables. This can be modeled as a Regression Problem. In such problems, the objective is to estimate the value of a continuous output variable.

4.3.1 Regression methods and Artificial Neural Networks

The linear regression consists of finding the best line that fits two variables so that knowing the value of one of them we can predict the value of the other, this limits us to work with only two variables. On the other hand, multiple linear regression is an extension of the linear regression in which more than two variables are involved and the data are fitted to a multidimensional surface [9]. However, the use of such methods would not be appropriate due to the low linear correlation between the climatic variables and the number of dengue cases, as could be observed in the analyzes performed in the previous section.

Artificial Neural Networks offer many potential advantages over alternative regression methods when dealing with problems that contain non-linear data and do not follow a normal distribution. Artificial Neural Networks are extremely versatile and do not require a formal specification of the model or the fulfillment of a certain probability distribution in the data [16]. Those are the main reasons why they were used in this work.

The error of a particular neural network configuration can be determined by running all training cases through the network and comparing the current output generated by the network to the expected output. The differences are combined by an error function, in this case the Root Mean Squared Error (RMSE) function.

4.3.2 Training and Validation

After the training of a neural network, the weights are adjusted for minimizing the error generated. Thus, the generated prediction results should be fairly close to the actual values considering the mentioned training set.

If the error of the trained network is relatively low, but the complete available dataset is used for training, it could happen that the network generates successful outputs for data included in that set, but when testing with new data, differing from those in the training dataset, unexpected outputs could be obtained, very different from those observed in the real world. This is because an overfitting has occurred with respect to the training and is not able to generalize for new cases. This can be solved by dividing the data set into two subsets: a training set used to adjust weights and a validation set used to perform independent checks on the progress of the learning algorithm.

As training progresses, also training and validation errors naturally decrease. However, if the validation error stops decreasing, and instead begins to increase, an overfitting is occurring, so the training must stop.

In the present work, the training dataset used corresponds to 70% of the total amount of data, and the validation dataset corresponds to the remaining 30%, chronologically sorted. This relation was selected after doing multiple tests with different values.

4.3.3 Data Preprocessing

Before introducing the collected data into the neural network for training, it should be noted that there are differences in the range and measure units of the variables. Variables related to Temperature (degrees centigrade), Humidity (percentage), rainfall (millimeters) and number of cases are considered. Using this data directly for the training of a neural network could cause a considerably delay in convergence. A solution to this problem is to normalize the training and validation data before using them. By normalizing this variables all data would be in the same range and measure unit.

In order to perform the normalization, it is necessary to know, firstly, the ranges of each of the variables to be used, that is, the maximum and minimum values of each variable. It is also necessary to know the range of values to which these variables will be mapped.

MIN-MAX [17] is the selected method to scale the values of the input variables of the neural networks to be trained in this work with a scale of 0 to 1. So, for example, for the temperature variable, if the min value is 3 degrees and the max value is 42 degrees, the map result value will be 0.3846.

First, the training set is normalized, and then, using the minimum and maximum values of the training set, the validation set is also normalized.

4.3.4 Neural Network Parameters Choice

There is an important issue when defining the architecture of the Neural Network to be used. The definition of parameters as the number of input and output neurons to use is relatively simple because they are defined by the problem, however there could be some uncertainty when defining the remaining parameters, such as: Number of hidden layers to use, Number of neurons per hidden layer and activation functions. Therefore, in this section we propose a method capable of evaluating different values for these parameters in order to find an optimal architecture to predict dengue cases.

The possible values used in the evaluation for the parameters are summarized in Table 2:

Table 2: Possible values for network parameters

Parameter	Value
Number of Hidden Layers	{1,2}
Nodes per Layer	{0, 1, ..., 2n}
Activation Functions	Log, Sigmoid, Tanh, ElliotSymmetric, Linear

Section 4.2 described the process to be followed to find the covariates with the highest correlation with the number of dengue cases, and thus generate at most 26 models. Then, for each of these models, test cases are generated, varying each of the parameters mentioned in Table 2.

For each test case a new neural network is constructed according to the corresponding parameters. The weights of the connections between nodes are initialized with random values, using a seed, that is stored for reproducibility and comparison purposes, to generate those values. Three different types of learning algorithms are used, so the test case is executed three times, generating combinations of different weight values, which produce different validation errors for the same architecture. The learning algorithms used to adjust the network weights are: ResilientPropagation, QuickPropagation and LevenbergMarquardt [17].

Once all the test cases are executed, the validation errors produced by each of the generated networks are analyzed, and the one with the least validation error is chosen as the best model.

Considering each district can obtain at most 26 models, and a potentially large number of test cases, is important to define the duration of the training process for each test case. Choosing a fixed number of iterations for the training of all the neural networks is very inefficient since for a significant number of cases the minimum validation error could be reached with a few iterations, performing the remaining iterations in vain and slowing down the general process of finding the best network architecture for a particular district. To avoid this, an Early Stopping strategy is used. This strategy consists in stopping the training if, after a minimum number of iterations, there is no significant improvement in the mean squared error of the network when calculated according to the validation set. As long as there are considerable improvements in the validation error of the network, training will continue until a maximum number of tolerated iterations is met. In the present work we used for the minimum number of tolerated iterations a value of 100, for the maximum number of tolerated iterations a value of 75,000 and 0.00001 as the minimum improvement of the validation error value.

4.3.5 Knowledge Discovery on neural networks

Artificial Neural Networks are generally considered to be "Black Boxes" because they are believed to provide little information about the contributions of input variables in the prediction process. However, there are several methodologies that can be used to estimate the influence of the input variables of a Neural Network. The method of Connection Weights is used, due to its good precision. This method calculates the product of the weights of the connections that exist between each of the input neurons and each one of the output neurons and then summarize the products for each input neurons, obtaining a value considered as the importance of that input variable. [10]

Another method for quickly and graphically visualize the importance of Neural Network input variables with simple architectures is to use Neural Interpretation Diagrams (NIDs). This type of diagrams basically consists in drawing the neural network with its corresponding nodes and connections, highlighting with thicker lines those connections that have associated a greater weight. It is expected that those variables that have a strong relation with the output variables have several thick connections between the layers. This qualitative interpretation can be challenging for large models, especially if the weights values change their signs after passing the hidden layers.

In this work, both the Connection Weights method and Neuronal Interpretation Diagrams to determine the impact of each of the input variables with respect to the output variable of the neural network are used. [18].

4.4 Developed Applications

In the previous sections the steps to find the most important variables in the prediction of dengue cases were explained. We also described the steps to find the appropriate architecture of the neural network to make predictions from these variables. This section describes the developed implementations for the use of this predictive tool. Two applications were developed, web and desktop, each one with different purposes.

4.4.1 Web Application

The Denguemaps application proposed by Ojeda et al [8] was extended by adding the following functionalities:

- Prediction heat map by district with up to four weeks of anticipation
- Visualization of the architecture of the neural network used for the prediction of cases of each district, including the weights of each connection and the summary of general importance of each co-variable using the Connection Weights algorithm mentioned before
- Interactive tool for dengue cases prediction requiring the introduction of necessary model data
- Visualization of historical data of the number of cases in each district
- Neural Network training and validation errors graph.
- Form to load new co-variables data to retrain the neural network in the future
- Automated prediction of number of cases for the next four weeks according to loaded data from last weeks.

4.4.2 Desktop Application

The web application mentioned is intended to use by the Ministry of Health of Paraguay, because the geographical visualizations include only the Paraguayan territory. The multiplatform desktop application was developed considering the possibility that any other country can make its own predictions, using the developed method, with its own data. This application includes all the functionalities included in the web application with the exception of geographical visualizations. In addition, being a desktop application, it uses the resources of the client machine to find the model and does not depends on any server in which the resources could be more scarce to be shared. Both applications facilitate the prediction as it is able to find the ideal model without making any previous analysis with the data. This is because the developed model is generic for any prediction of dengue cases regardless of geographic region, only taking into account the climatic variables and the historical number of cases in the area to be predicted.

5. Evaluation and Experimental Results

This section describes the results obtained in the prediction of dengue cases for the different districts.

5.1 Tools and Test Environment

The Java programming language was used, with the framework for Machine Learning Encog [17], in computers with the following characteristics:

- Processor: Intel(R) Core(TM) i7-4510U
- CPU: 2.00GHz, 4 cores

- RAM Memory: 12 GB DDR3 1600 MHz
- Operating System: Ubuntu 14.04 64 bits

5.1.1 Experimental Results

In the table 5 it can be observed the model used for Asuncion and the predicted weeks of anticipation, from 1 to 4.

The lag column represents the anticipation for prediction, and the variables column shows the variables used. The number next to each variable represents the weeks before of the predicted variable, where 1 indicates the current week, 2 the previous one and so on. So, for example, for predicting the number of cases in Asuncion one week ahead, the following variables were used: number of cases of current week, the average temperature 11 weeks ago, rain accumulated 3 weeks ago and the average minimum humidity 12 weeks ago. It should be noted that all models include the number of cases as co-variable, since the presence or absence of the disease is determinant to perform the prediction of number of cases. This is because, although climatic conditions are ideal for the appearance of the transmitter mosquito, if the disease is not present, it will not spread. It can also be noted that, as explained in section 4.2, the same type of variable does not appear twice since the correlation between them is high and only generates overinformation to the model. In addition it can be noted that for all cases the model with at least one climatic variable was always selected, thus demonstrating its influence on the number of cases. The table 6 shows the errors obtained for Asuncion using the Root mean squared error method, previously mentioned. It may be noted that it is relatively low, considering a population number of 525,294. Figure 1 shows the predicted cases versus the real ones. In addition, a comparison of the results obtained using a linear regression was performed. In all cases the error thrown by the Neural Network was smaller. Also, can be notice how the error increases as anticipation number of weeks increase. For example, for Asunción, this evolution is shown in the Figure 2. Although the error in week 4 increased significantly over week 1, this prediction serves to get a rough idea of the number of cases that will occur within a month. This number will become increasingly accurate as the weeks pass and new data are available.

In addition, the ideal network architectures obtained for Asunción for each lag can be observed in Table 4, as well as the importance of each variable of the first model according to the Connection Weigths method showed in the Table 3

For the other 13 districts similar results were obtained and these can be visualized in the developed web application, available in `http://dengue.cds.com.py`

6. Conclusions and Future Works

The lack of tools to predict and quantify the spread of a disease is a major difficulty for estimating resources and actions needed by institutions responsible for ensuring the health of the population. In the present work, the number of dengue cases per district in Paraguay could be predicted with an anticipation up to 4 weeks, with errors proportional to the number of anticipation weeks of the prediction. In addition, it was possible to identify the climatic variables that affect the increase or decrease of the number of cases, through the analysis of the results thrown by the neural network. It was possible to verify how the different climatic variables affect the number of cases for the different districts. For example, for Asuncion the climatic variables that had the most influence were: the temperature with 11 weeks of anticipation, the accumulated rain with 3 weeks of anticipation and the relative minimum humidity with 12 weeks of anticipation. It can be noticed that temperature and humidity influence with a large delay in the number of cases and this finding is not presumed at first sight. Although there are other potentially usable variables (such as demographic, movement of people,

Table 3: Variables Importance for predicting cases in Asunción

Variable	Importance
Number of cases -1	-10,96
Maximum Temperature -10	-12,40
Rain -2	16,12
Minimum Humidity -11	27,33

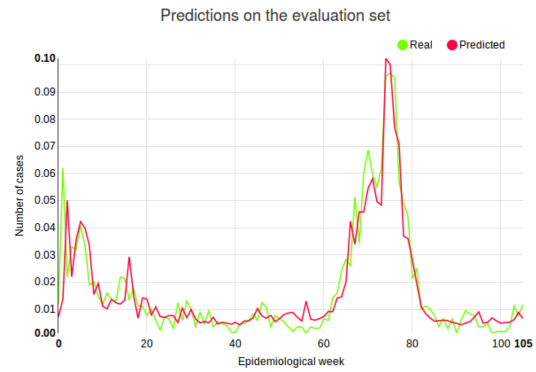

Fig. 1: Real vs Predicted cases for Asunción

Table 4: Network Architecture used for the prediction of cases for Asunción

Lag	Nodess Layer	Activation
1	4-5-5-1	Tanh, Sigmoid
2	3-2-6-1	Log, Sigmoid
3	4-8-1	Sigmoid, Sigmoid
4	3-5-1	Sigmoid, Sigmoid

Table 5: Resultant models for Asuncion

Lag	Variables
1	cases1, max_temperature11, rain3, min_humidity12
2	cases2, max_temperature11, min_humidity12
3	cases3, max_temperature11, lluvia3, min_humidity12
4	cases4, max_temperature12, avg_humidity12

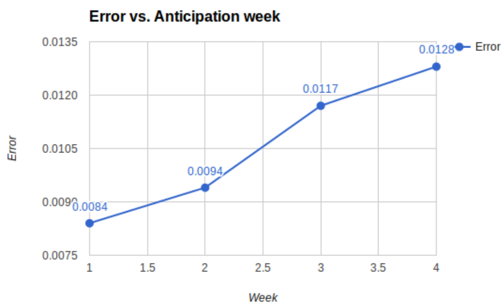

Fig. 2: Error evolution according to the weeks of anticipation for Asunción

Table 6: Error, R and error with the MLR model for Asuncion

Lag	Error	R	MLR
1	14.157	0.9164	135.11
2	15.850	0.8988	139.40
3	19.731	0.8458	129.35
4	21.568	0.7926	138.46

activities against dengue, larval infestation) the quality of them in our country, especially for inland regions, are low or are added with a very high level and its use would only generate noise, reason why they were not used in this work. We compared the results obtained by the Artificial Neural Network against a linear model, demonstrating that the Neural Network obtains more accurate results for all cases, thus proving that the climatic variables are non-linearly related to the number of dengue cases. A method of selecting climatic variables was developed that could potentially be used in any geographic region without need to perform previous analyzes on the data, thus facilitating prediction. The methods developed in this work were added to the existing web application developed by Ojeda et al. [8] which automatically performs predictions with data uploaded once a week, and allows re-training of the neural network with new data when necessary. In addition, a desktop application was developed to allow the use of the tool with data that are not from Paraguay.

Also, it was noticed that the model change for each district, since the conditions are different, reason why the training and prediction was realized for each of them individually, avoiding to generate noise in the prediction of others.

This is the first work in Paraguay that demonstrates the relationship between number of cases and climate. All the results obtained can be visualized in the web application `http://dengue.cds.com.py`

6.1 Future Works

As future work, the automation of the collection of data, both climatic and number of cases can be proposed, so predictions can be performed fully automatically without having to load data manually. More types of variables can be incorporated, not only climatic, to the prediction, for example, demographic, movement of people, activities against dengue, larval infestation among others. If the data exists, it could perform the prediction on a daily and non-weekly basis, in order to perform a prediction with a lower granularity. Also, other prediction methods can be implemented, and compared with this proposal.

References

[1] World Health Organization. Official site [online] Available on: http://www.who.int/mediacentre/factsheets/fs117/es/
[2] World Health Organization. Official site (2016, Setiembre 29) [online] Available on: http://www.who.int/mediacentre/factsheets/fs387/es/index2.html
[3] Paraguayan Health Surveillance Office. Boletín Epidemiológico Semanal Edición Número 49, Semana Epidemiológica: 51, 26 de diciembre del 2013 [online] Available on: http://vigisalud.gov.py/boletines/25_08_2016_12_46_10_2013.12.30-boletin-epidemiologico.pdf
[4] B. Gultekin, M. Sari y H. M. Aburas. Reconigtion of Dengue Disease Patterns Using Artificial Neural Networks. 5 th International Advanced Technologies Symposium, Mayo, 2009, Karabuk, Turquía
[5] P.H.M. Nishanthi Herath, A.A.I. Perera, H.P.Wijekoon. Prediction of Dengue Outbreaks in Sri Lanka using Artificial Neural Networks. International Journal of Computer Applications (0975 – 8887) Volumen 101– No.15, Septiembre 2014
[6] Y. Ling Hii. Climate and Dengue Fever: Early Warning based on temperature and rainfall. Umea University, Sweden. 2013
[7] M. Báez. Modelo predictivo de focos de dengue aplicado a Sistemas de información geográfica. Universidad Nacional de Asunción. 2014
[8] V. Ojeda, N. Valdez, J. Paciello, J. Pane. Estandarización de Reporte de Casos y Predicción de Brotes de Dengue en Paraguay en Base a Datos Abiertos. Universidad Nacional de Asunción. 2016
[9] J. Han, M. Kamber. Data Mining, Concepts and Techniques. Morgan Kaufmann, Second Edition. 2006
[10] Olden, Joy, Death. An accurate comparison of methods for quantifying variable importance in artificial neural networks using simulated data. 2004
[11] World Health Organization (WHO). The world health report 2002: Reducing risks and promoting healthy life. Geneva - 2002. (2016, Setiembre 29) [online] Available on: http://www.who.int/whr/2002/en/whr02_en.pdf
[12] G. L. Rúa-Uribe, C. Suárez-Acosta, J. Chauca, P. Ventosilla, R. Almanza. Modelado del efecto de la variabilidad climática local sobre la transmisión de dengue en Medellín mediante análisis de series temporales. Universidad de Antioquia, Colombia. 2013
[13] C. D. Manning, P. Raghavan and H. Schütze. Introduction to Information Retrieval, Cambridge University Press. 2008 [online] Available on http://nlp.stanford.edu/IR-book/html/htmledition/feature-selection-1.html
[14] C. W. Morin, A. C. Comrie y K. Ernst. Climate and Dengue Transmission: Evidence and Implications. University of Arizona, School of Geography and Development. 2013
[15] R. May, G. Dandy and H. Maier. Review of Input Variable Selection Methods for Artificial Neural Networks, Artificial Neural Networks - Methodological Advances and Biomedical Applications, InTech, DOI: 10.5772/16004. 2011
[16] J. Montaño, A. Palmer, P. Muñoz. Artificial neural networks applied to forecasting time series. Universidad de las Islas Baleares y Universidad Politécnica de Cataluña. 2010
[17] J. Heaton. Programming Neural Networks with Encog3 in Java. Heaton Research, Inc. 2011
[18] beckmw. Variable importance in neural networks [Blog post]. 2013. Available on: https://beckmw.wordpress.com/2013/08/12/variable-importance-in-neural-networks/