Prediction of Dengue Cases in Paraguay Using Artificial Neural Networks

Yohanna Lisnichuk <ylisnichuk@pol.una.py>

Víctor Ughelli <vughelli@pol.una.py>

Julio Paciello <julio.paciello@pol.una.py>

Juan Pane <jpane@pol.una.py>

July, 2017

Content

- Introduction
- Problem Definition
- Related Works
- Proposed Solution
- Experimental Results
- Conclusions
- Future Works

Introduction

Introduction

Problem Definition

Related Works

Proposed Solution

Experimental Results

Conclusions

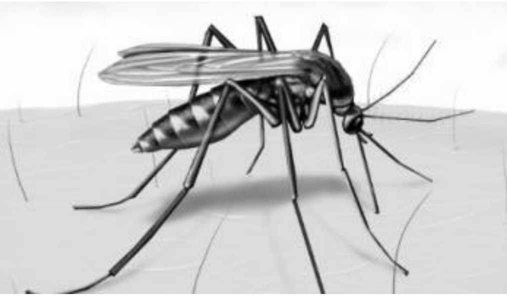
Future Works

Dengue Fever

Dengue Fever is a viral disease transmitted by mosquitoes that has spread rapidly in recent years in all regions[1].

Currently, **about 40%** of the world's population is at risk to get dengue fever

This disease **can be fatal** and has already claimed thousands of lives worldwide.



[1] World Health Organization website, Available on: http://www.who.int/mediacentre/factsheets/fs117

Dengue Fever

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Macro-factors that influence the increase in the number of dengue cases:

- Climate change
- Global warming
- "El Niño" phenomenon

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Case Study: Paraguay

- Population: around 6.5million
- Capital: Asunción
- Overall climate:
  tropical to subtropical
  (ideal for reproduction of Aedes
  aegypti).

Introduction

Problem Definition

Related Works

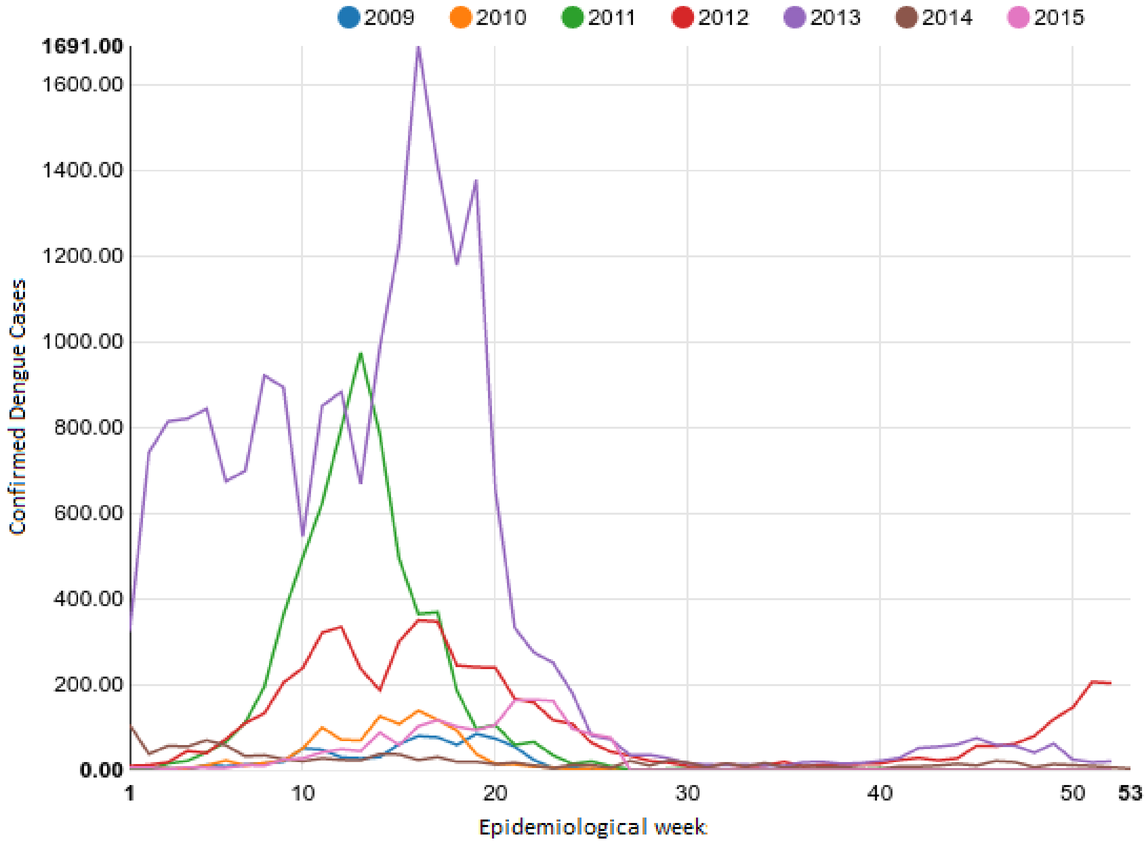
Proposed Solution

Experimental Results

Conclusions

Future Works

Dengue Fever in Paraguay



Weekly Confirmed Dengue Cases in Asunción from week 1- 2009 to week 52 - 2015

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Prediction of Dengue Cases

The prediction of dengue cases with some anticipation could help to provide:

- Human Resources
- Financial Resources
- Hospital and Medical Centres

Before a peak of cases occurs

Problem Definition

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Problem Definition

- Lack **of a predictive model** of the number of dengue cases with a **reasonable anticipation** time that will allow the authorities in Paraguay to make better decisions in a timely manner.

Also:

- The climate co-variables with the highest correlation respect to the number of dengue cases are not identified in the study area.

- Lack of an application that integrates a predictive model of the number of dengue cases in the study area.

Related Works

Global Level

- Gultekin et al. [1] – Singapore

  average temperature, relative humidity, rain per week

- Nishanthi et al. [2] – Sri Lanka

  temperature and humidity (1 and 4 weeks lags) rainfall (l4 weeks lag) and number of cases from the previous week

- Ling Hii [3] – Singapore

  average temperature, rainfall (9 to 16 weeks lags)

- Rua-Uribe [4] – Colombia

  rainfall (rainfall 20 weeks)

[1] B. Gultekin Cetiner, Murat Sari y Hani M. Aburas. Reconigtion of Dengue Disease Patterns Using Artificial Neural Networks. 5 th International Advanced Technologies Symposium (IATS'09), Mayo 13-15, 2009, Karabuk, Turquía

[2] 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)
[3] Yien Ling Hii. Climate and Dengue Fever: Early Warning based on temperature and rainfall. Umea University, Sweden. 2013

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Paraguay

- Báez [1]
- Ojeda et al. [2]
  - Outbreak classifier
  - 1 week in advance
  - No related co-variables identification
  - Web application

[1] Maximiliano Báez. Modelo predictivo de focos de dengue aplicado a Sistemas de información geográfica. Facultad Politécnica, Universidad Nacional de Asunción. 2014

[2] Verena Ojeda, Natalia Valdez, Julio Paciello, Juan Pane. Estandarización de Reporte de Casos y Predicción de Brotes de Dengue en Paraguay en Base a Datos Abiertos. Facultad Politécnica, Universidad Nacional de Asunción. 2016

Proposed Solution

Proposed Solution

- Introduction
- Problem Definition
- Related Works
- Proposed Solution
- Experimental Results
- Conclusions
- Future Works

- Predictive model of number of dengue cases using Artificial Neural Networks
- Identification of co-variables that affects the number of dengue cases
- Integration of the predictive model within web/desktop applications.

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

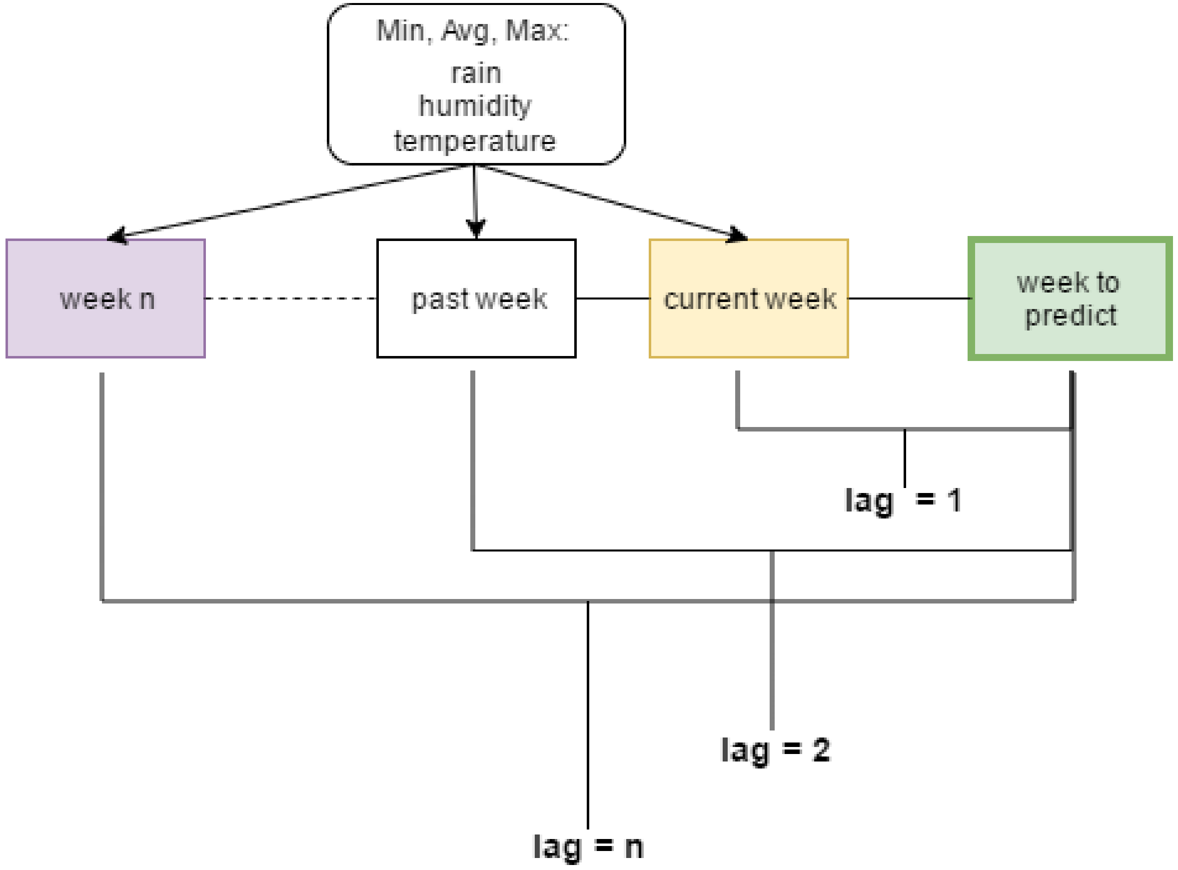
Data Collection

- Number of Dengue cases
- Climatic variables: rain, temperature, humidity
- River levels

http://vigisalud.gov.py/boletin_epidemiologico
https://www.wunderground.com/
http://meteorologia.gov.py/nivel/

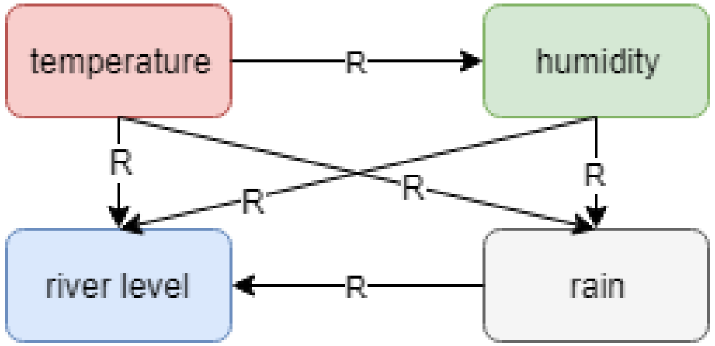
Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Feature Selection



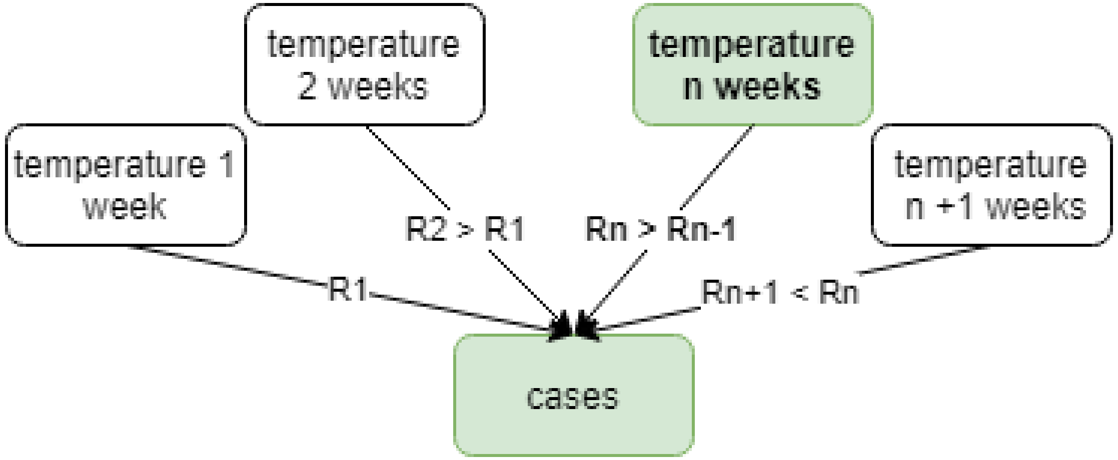
Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Feature Selection



1) Pearson Correlation between each variable

2) Pearson Correlation between each variable and the dengue cases variable



Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Feature Selection

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

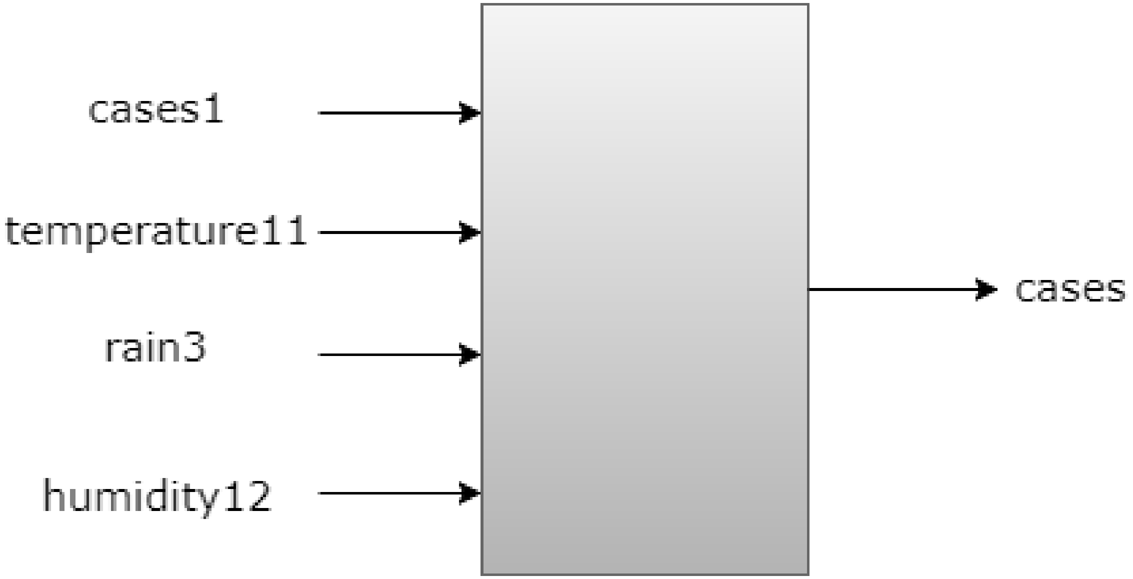
```
for Variable (humidity, temperature, rain) delay 1 ... n do
    variables += Find two better variables with higher Pearson correlation with number of cases
end for
combinations = combine(variables,number of cases of the previous week) (number of cases, humidity1, humidity2, rain1, rain2, temperature1, temperature2)
return combinations
```

Feature Selection

Model	Variables Combination
1	casesVarA
2	casesVarA, temperatureVarA
3	casesVarA, temperatureVarB
4	casesVarA, humidityVarA
…	
27	casesVarA, temperatureVarB, humidityVarB, rainVarB

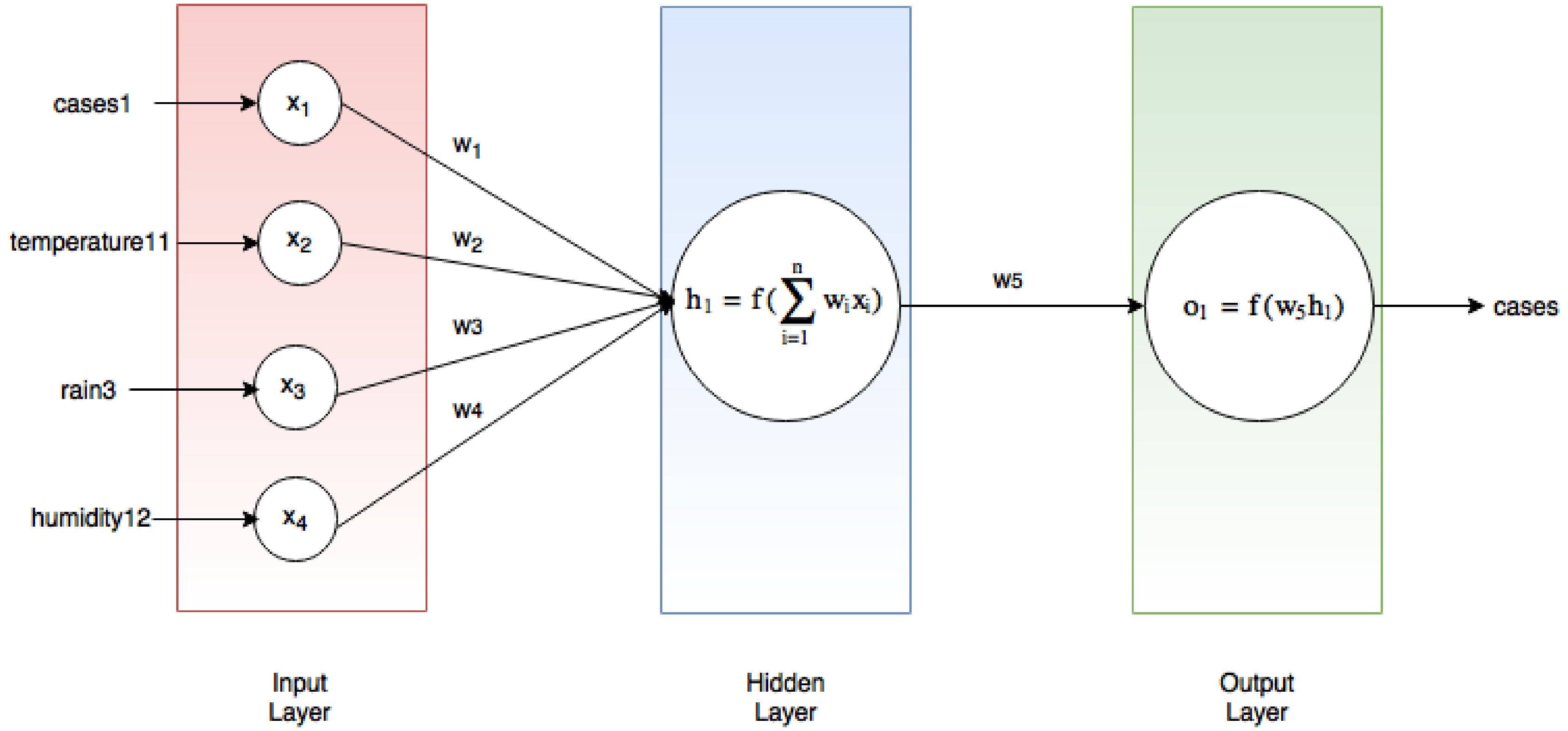
Predictive model Implementation

Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works



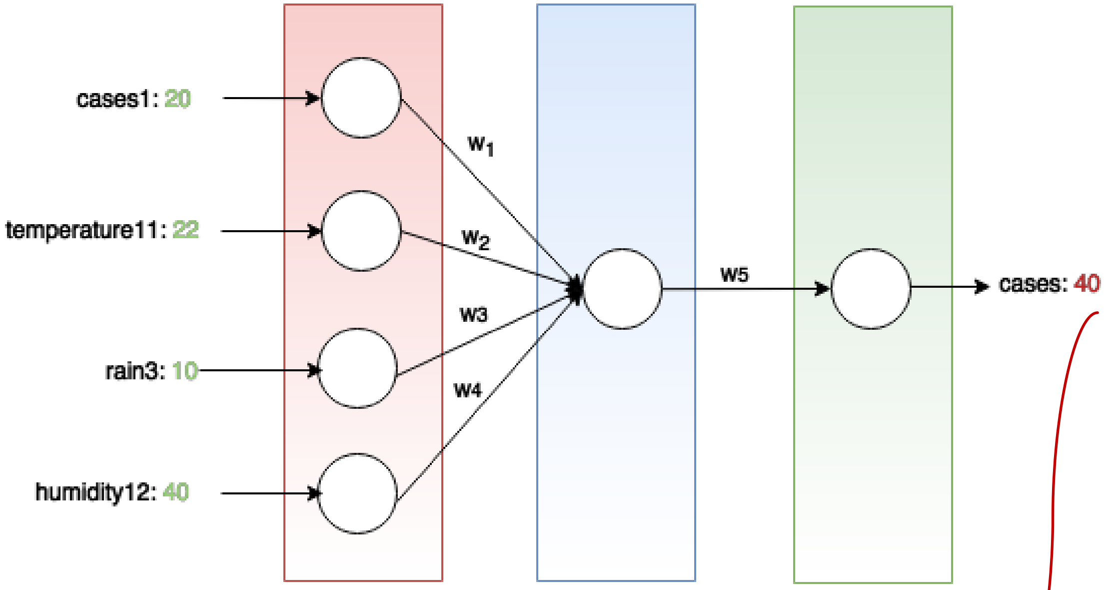
Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Artificial Neural Networks



Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Artificial Neural Networks



(Predicted)

Training Dataset

cases1	Temperature 11	rain3	humidity11	cases
20	22	10	40	**35**
35	20	5	20	20
…	…	…	…	…

(Expected)

Introduction

Problem Definition

Related Works

Proposed Solution

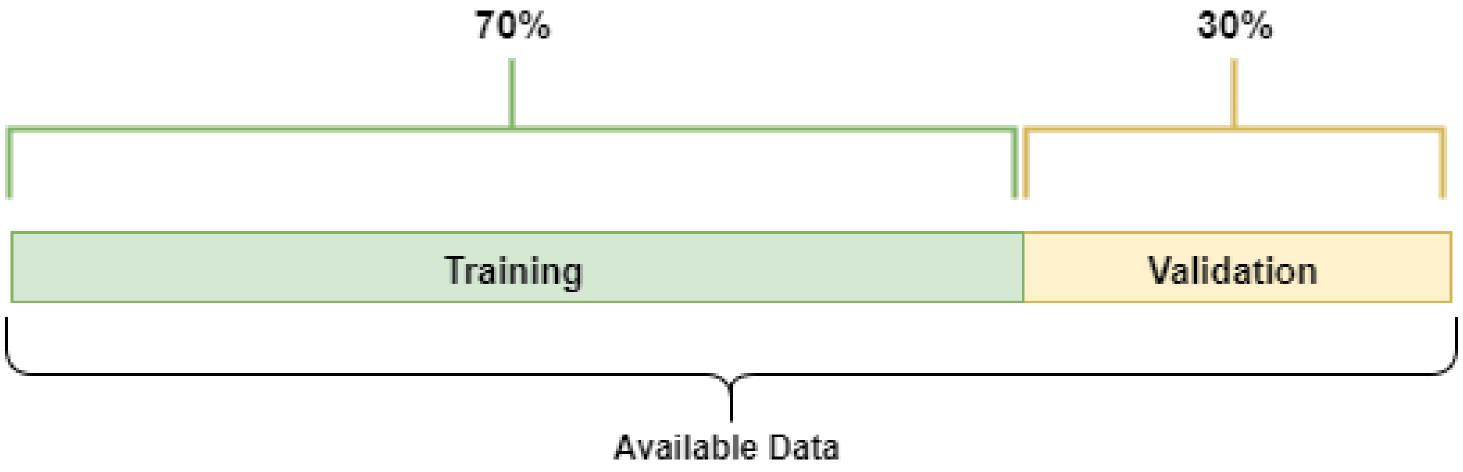
Experimental Results

Conclusions

Future Works

Implementation Details

- Error function: RMSE[1]
- Data pre-processing: MINMAX
- Training and Validation datasets



[1]Root Mean Squared Error

Implementation Details

Introduction

Problem Definition

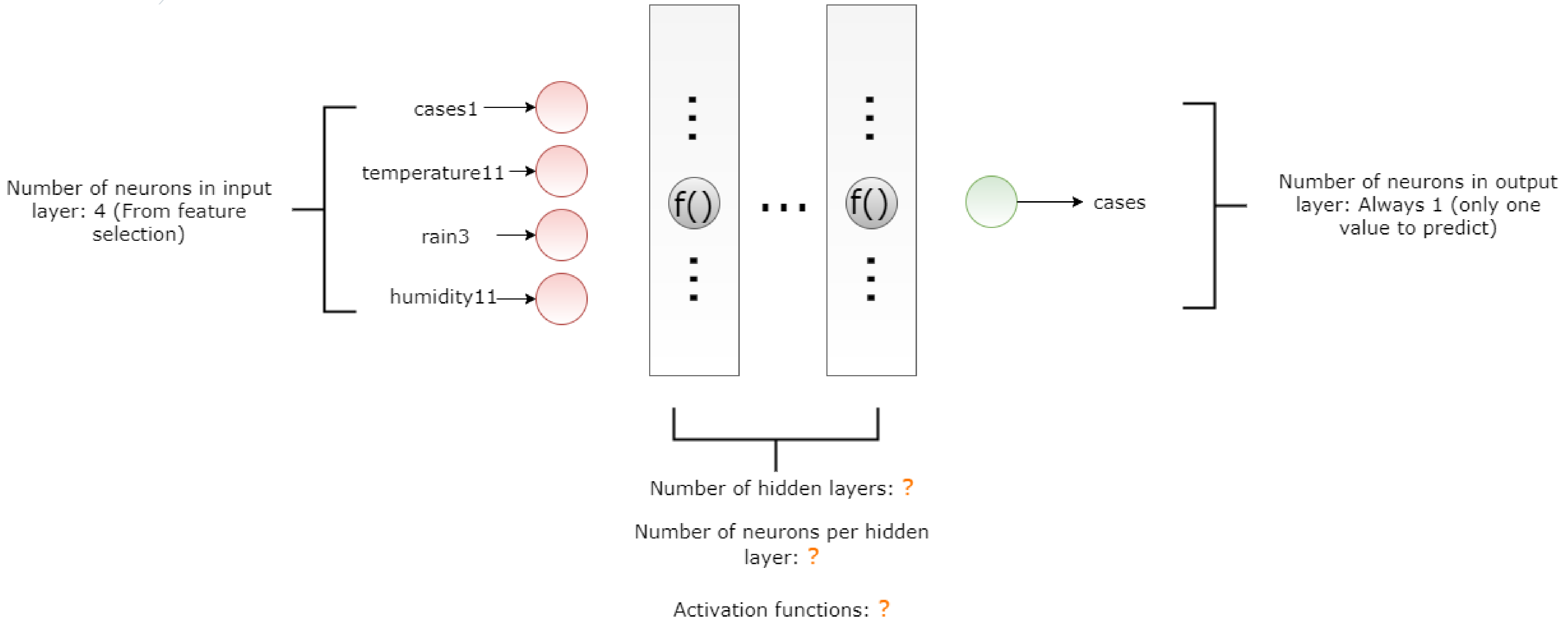
Related Works

Proposed Solution

Experimental Results

Conclusions

Future Works



Introduction
Problem Definition
Related Works
Proposed Solution
Experimental Results
Conclusions
Future Works

Implementation Details

- Parameters Choice

Parameter	Value
Number of Hidden Layers	{1, 2}
Number of Neurons per Hidden Layer	{0, 1, …, 2n}
Activation Functions	Log, Sigmoid, Tanh, Elliot Symmetric, Linear

Implementation Details

For a specific District and Lag the following test cases are analyzed:

Variables	Nodes per hidden layer	Activation functions	Error
Cases1, temperature11,rain3, humidity11	1,1	Sigmoid, Sigmoid	0.0014
Cases1, temperature11,rain3, humidity11	1,1	Sigmoid, Tanh	0.0025
…	…	…	…
Cases1, temperature11,rain3, humidity11	1,2	Sigmoid, Sigmoid	0.0050
…	…	…	…
Cases1, temperature2, rain2, humidity2	1,1	Sigmoid, Sigmoid	0.0070
…	…	…	…

Web Application

Introduction

Problem Definition

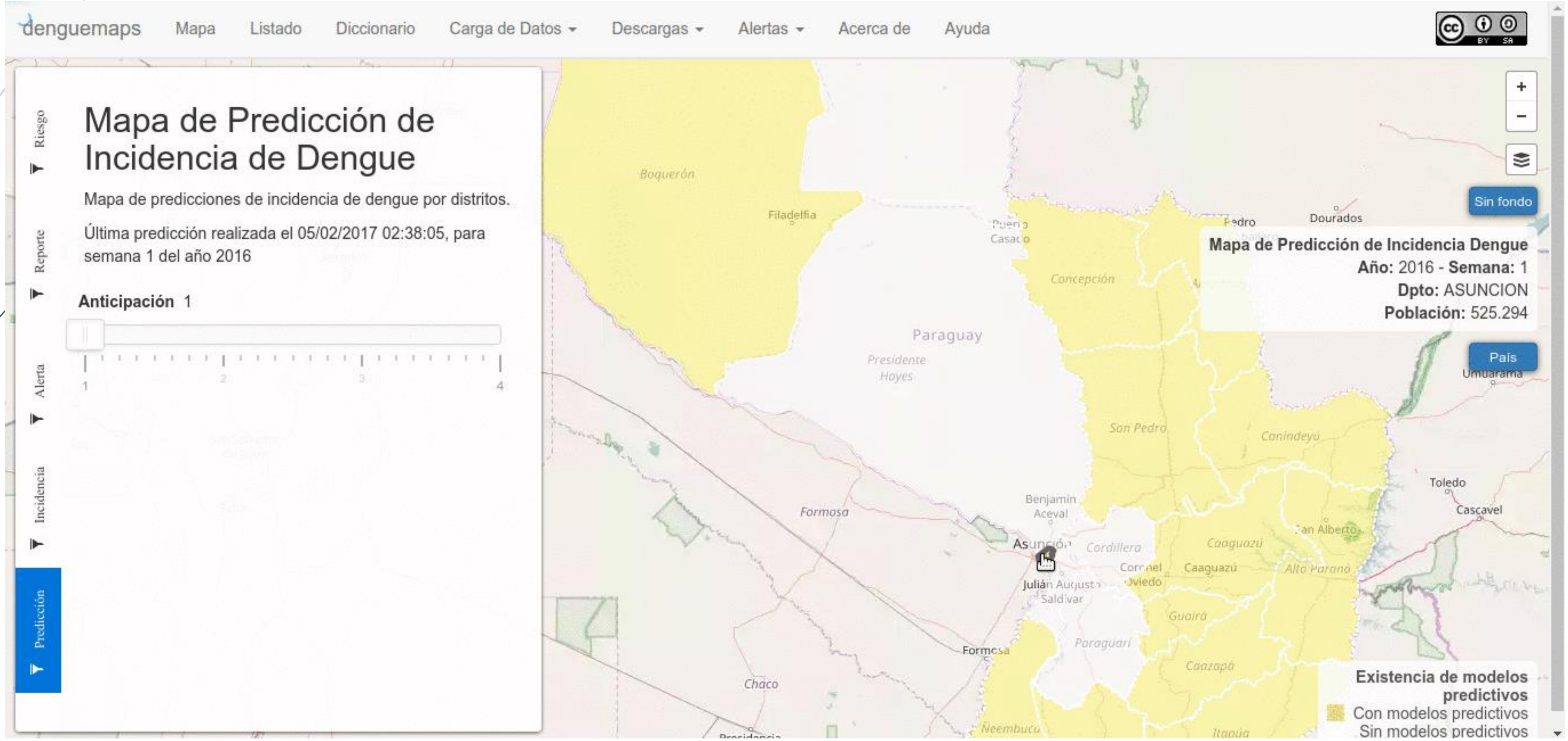
Related Works

Proposed Solution

Experimental Results

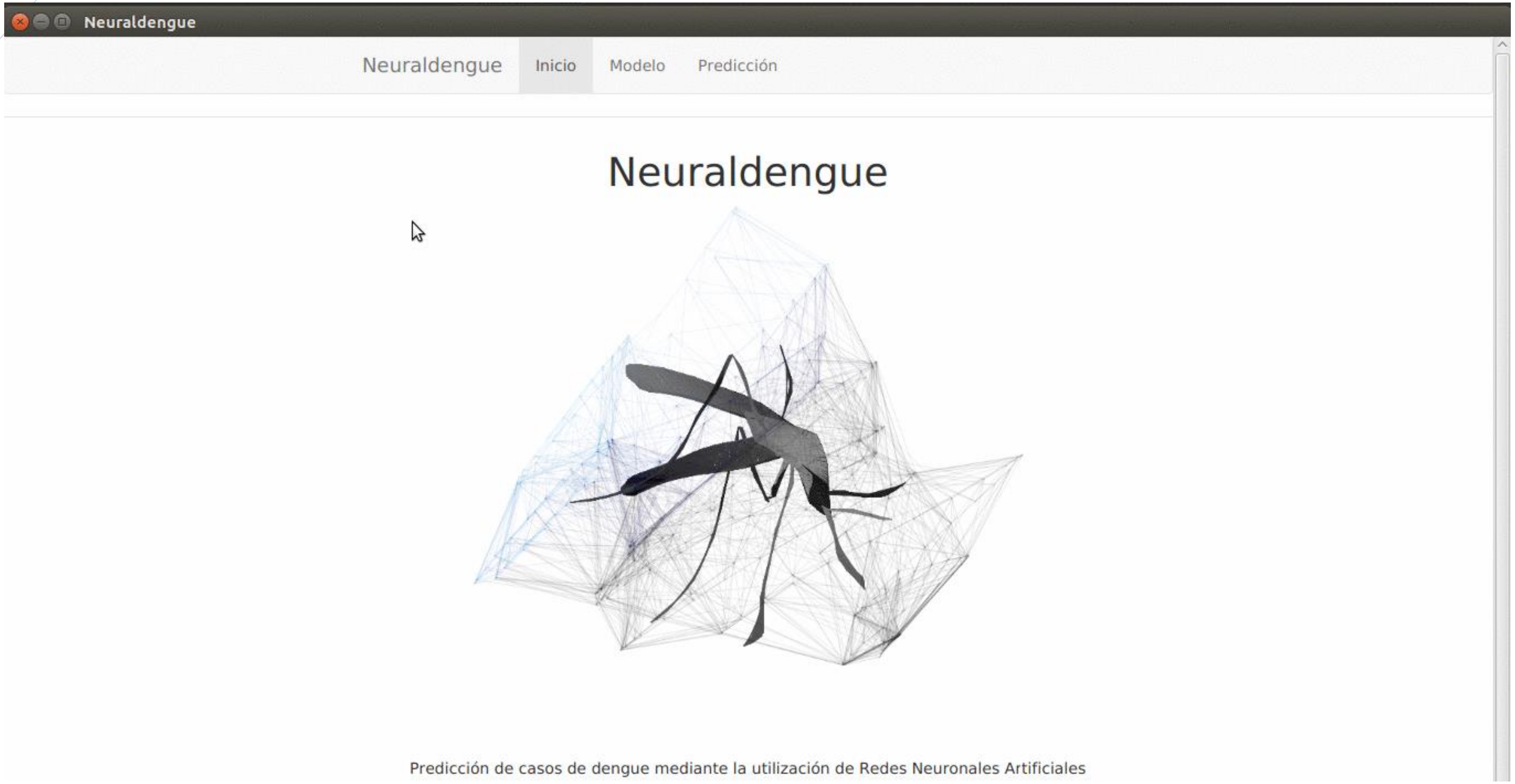
Conclusions

Future Works



- Introduction
- Problem Definition
- Related Works
- Proposed Solution
- Experimental Results
- Conclusions
- Future Works

Desktop Application



Experimental Results

Introduction

Problem Definition

Related Works

Proposed Solution

Experimental Results

Conclusions

Future Works

Evaluation Methodology

- Prediction with data from 14 districts
- 366 records per district (from week 1 - 2009 to week 52 - 2015)
- First 256 weeks for training (70%) and last 109 for validation (30%)
- At most 27 models tested per district
- RMSE to measure the error

Resultant models

Lag	Variables	Neurons per Layer	Activations
1	cases1, max_temperature11, rain3, min_humidity12	Input: 4 Hidden1: 5 Hidden2: 5 Output: 1	Hidden: Tanh Output: Sigmoid
2	cases2, max_temperature11, min_humidity12	Input: 3 Hidden1: 2 Hidden2: 6 Output: 1	Hidden: Log Output: Sigmoid
3	cases3, max_temperature11, rain3, min_humidity12	Input: 4 Hidden1: 8 Output: 1	Hidden: Sigmoid Output: Sigmoid
4	cases4, max_temperature12, avg_humidity12	Input: 3 Hidden1: 5 Output: 1	Hidden: Sigmoid Output: Sigmoid

Resultant models for Asunción

Prediction Accuracy

Introduction

Problem Definition

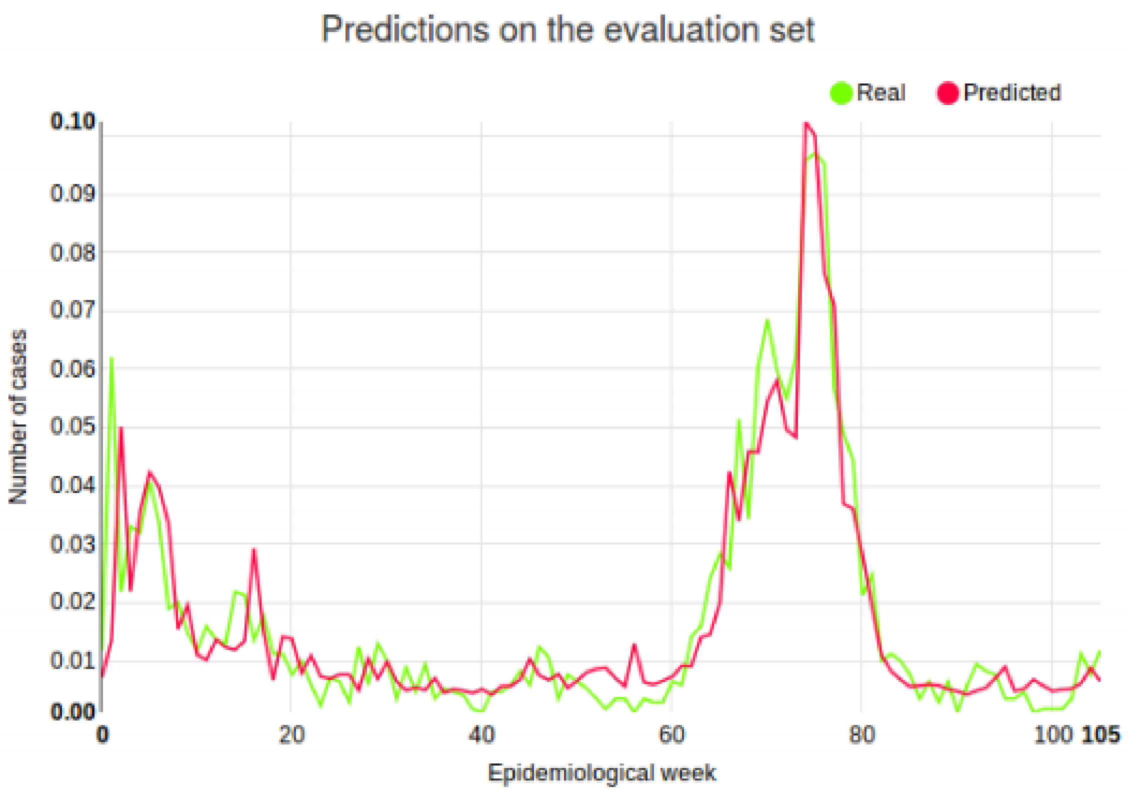
Related Works

Proposed Solution

Experimental Results

Conclusions

Future Works



Real vs Predicted cases for Asunción

Conclusions

Introduction

Problem Definition

Related Works

Proposed Solution

Experimental Results

Conclusions

Future Works

Conclusions

The contributions of this work are:

- Prediction of number of dengue cases per district in Paraguay with an anticipation up to 4 weeks
- Identify the climatic variables that affect the number of cases in Paraguay
- A method for selecting climatic variables for prediction of dengue cases
- Web and desktop applications

Future Works

Introduction

Problem Definition

Related Works

Proposed Solution

Experimental Results

Conclusions

Future Works

Future Works

- Automation of the collection of data
- More types of variables can be incorporated to the prediction, not only climatic.
- Perform the prediction on a daily and non weekly basis
- Other prediction methods can be implemented, and compared with this proposal.


THANK
YOU
GRACIAS
ARIGATO
SHUKURIA
DANKSCHEEN
TASHAKKUR ATU
YAQHANYELAY
SUKSAMA
EKHMET
BIYAN
SHUKRIA
TINGKI
GRAZIE
MEHRBANI
MAAKE
KOMAPSUMNIDA
GOZAIMASHITA
EFCHARISTO
JUSPAXAR
PALDIES
BOLZÏN
MERCI