DengAI: Predicting Disease Spread

Group C : G.A Akila Shashith Ganegoda, M.L.N.Kaushalee, K. G. Inuka Manmitha, R Nilakshan, S.P.A.G.T Rajapaksha

Introduction

Using environmental data collected by U.S. Federal Government agencies, can you predict the number of dengue fever cases reported each week in San Juan, Puerto Rico and Iquitos, Peru?

- Dengue fever is a mosquito-borne disease in tropical and subtropical parts of the world.
- Symptoms of dengue fever can range from mild to severe, and in severe cases can lead to death.
- Climate change will likely produce distributional shifts with significant public health implications worldwide.
- Dengue fever has spread in recent years, with many cases occurring in Latin America.
- Predicting the number of dengue cases weekly in specific locations can help improve research initiatives and resource allocation to help fight life-threatening pandemics.

Problem Definition

Using environmental data to aid in public health initiatives, predict the number of dengue fever cases reported weekly in San Juan, Puerto Rico, and Iquitos, Peru.

Description of the Dataset

- The goal is to predict total cases for each (city, year, weekofyear) in the test set.
- The test set has data for two cities, San Juan and Iquitos, spanning 5 and 3 years, respectively.
- The test data are sequential and non-overlapping with any training data.
- The features include climate data, precipitation measurements, dew point temperature, air temperature, relative humidity, specific humidity, and vegetation index.

- city – City abbreviations: sj for San Juan and iq for Iquitos
- week_start_date – Date given in yyyy-mm-dd format
- station_max_temp_c – Maximum temperature
- station_min_temp_c – Minimum temperature
- station_avg_temp_c – Average temperature
- station_precip_mm – Total precipitation
- station_diur_temp_rng_c – Diurnal temperature range
- precipitation_amt_mm – Total precipitation
- ndvi_se – Pixel southeast of city centroid
- ndvi_sw – Pixel southwest of city centroid
- ndvi_ne – Pixel northeast of city centroid
- ndvi_nw – Pixel northwest of city centroid

- reanalysis_sat_precip_amt_mm – Total precipitation
- reanalysis_dew_point_temp_k – Mean dew point temperature
- reanalysis_air_temp_k – Mean air temperature
- reanalysis_relative_humidity_percent – Mean relative humidity
- reanalysis_specific_humidity_g_per_kg – Mean specific humidity
- reanalysis_precip_amt_kg_per_m2 – Total precipitation
- reanalysis_max_air_temp_k – Maximum air temperature
- reanalysis_min_air_temp_k – Minimum air temperature
- reanalysis_avg_temp_k – Average air temperature
- reanalysis_tdtr_k – Diurnal temperature range

Proposed Solution

We propose to utilise the dataset for a predictive task, such as a regression task, to forecast the total cases of dengue fever predictions (city, year, week of year) and the predicted total cases.

Mapping the Problem: The solution involves mapping environmental factors to disease incidence, aiding resource allocation and research efforts to combat pandemics.

(Resource: https://www.drivendata.org/competitions/44/dengai-predicting-disease-spread/page/80/)