

MODELO PREDICTIVO PARA LA ESTIMACIÓN DE LOS NÚMEROS DE CASOS DE DENGUE EN LOS DEPARTAMENTOS DEL PERÚ

Christopher Ballon, Hernán Hernández, Ruben Neyra, Christian Ballon

RESUMEN

Se implementó un modelo predictivo para estimar el número de casos de dengue por región considerando datos climáticos.

OBJETIVO

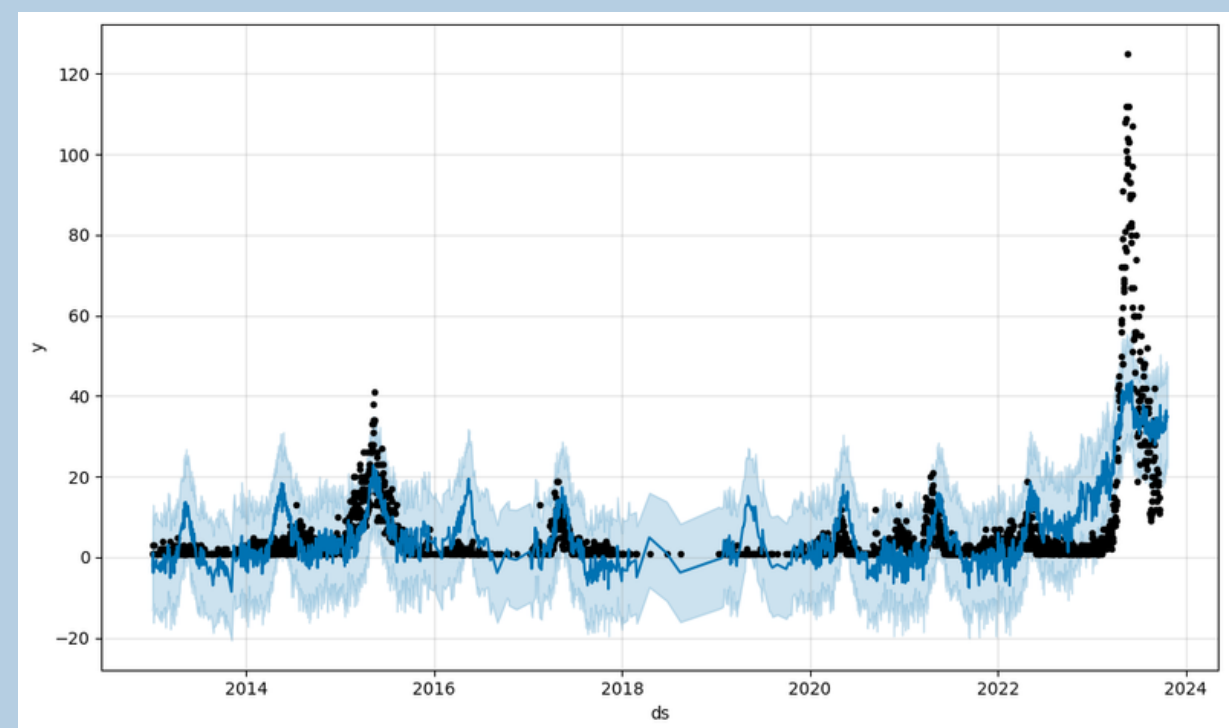
Predecir el número de casos de dengue por región del Perú considerando variables climáticas y demográficas

METODOLOGÍA

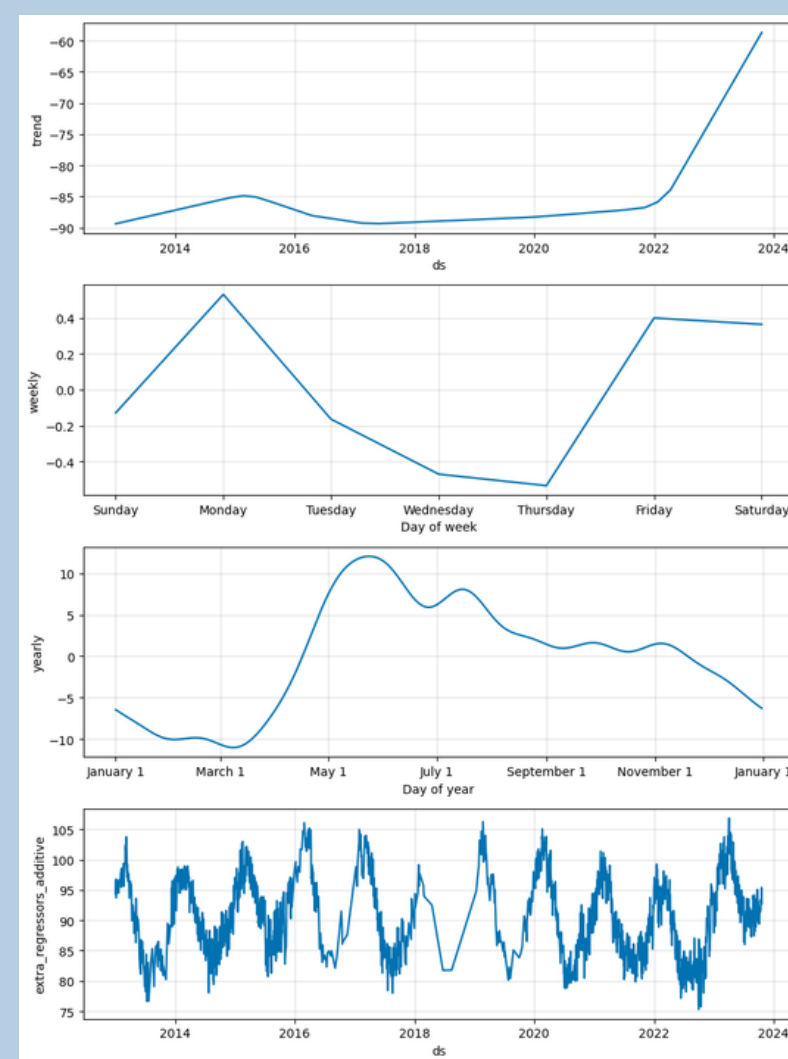
El modelo elegido fue de tipo supervisado, concretamente utilizamos un modelo predictivo basado en análisis de series temporales, realizamos el entrenamiento con datos de número de casos confirmados por fecha en cada provincia del país, debido a su elevada densidad poblacional, tomamos a Lima en regiones Norte, Sur, Este y Oeste y como conjunto de validación el número de casos de los últimos 30 días, además consideramos como variables regresoras para el modelo, la temperatura máxima de la región en cuestión y precipitación. Realizamos limpieza y preprocesado de datos cómo eliminar caracteres inválidos, esto nos permitió realizar la integración con las tablas maestras que nos proporcionaron correspondiente a los datos de centros de salud, también con datos meteorológicos obtenidos para cada región.

RESULTADOS

Datos históricos con el número de casos (puntos en negro) y predicción (curva en azul) considerando el número de casos y las variables temperatura máxima y precipitación



Análisis por componentes



MODELO CONSIDERANDO NÚMERO DE CASOS, MÁS LAS VARIABLES REGRESORAS: TEMPERATURA MÁXIMA Y PRECIPITACIÓN:
MAE 27.56 (DIFERENCIA PROMEDIO DE NÚMERO DE CASOS PREDICHOS Y VERDADEROS)
MAPE 8.04 (DIFERENCIA PORCENTUAL DEL PROMEDIO DE NÚMERO DE CASOS PREDICHOS Y VERDADEROS)

UTILIDAD DE LA PROPUESTA

El modelo una vez validado y optimizado permitirá integrarse a sistemas existentes en el MINSA para la predicción del número de casos de dengue.

DATASETS UTILIZADOS

Datasets proporcionados durante el datathon y otros que contienen datos climáticos