

Una revisión sistemática de Modelos de clasificación de dengue utilizando machine learning

Gisella Luisa Elena Maquen-Niño¹, Jessie Bravo², Roger Alarcón³,
Ivan Adrianzén-Olano⁴, Hugo Vega-Huerta⁵

gmaquenn@unprg.edu.pe; jbravo@unprg.edu.pe; ralarcong@unprg.edu.pe;
ivan.adrianzen@untrm.edu.pe; hvegah@unmsm.edu.pe

^{1,2,3} Grupo de Investigación en transformación digital, Universidad Nacional Pedro Ruiz Gallo, Calle Juan XXIII 391, C.P. 14001, Lambayeque, Perú

⁴ Universidad Nacional Toribio Rodríguez de Mendoza, Calle Higos Urco N° 342, C.P. 01000, Chachapoyas, Perú

⁵ Universidad Nacional Mayor de San Marcos UNMSM, Lima, Perú

DOI: [10.17013/risti.50.5-27](https://doi.org/10.17013/risti.50.5-27)

Resumen: El dengue es una enfermedad arboviral que anualmente reporta un gran número de infectados en la costa norte y la selva peruana. Según las estadísticas, está aumentando cada año. Este artículo tiene como objetivo desarrollar una revisión sistemática de la literatura científica sobre las variables de estudio y los métodos de aprendizaje automático utilizados actualmente para detectar la infección por dengue. La metodología utilizada fue PRISMA, mapeando inicialmente la literatura de 274 artículos científicos, quedando seleccionados 33 artículos para la revisión sistemática. Los resultados obtenidos son que los algoritmos de aprendizaje automático más utilizados son las redes neuronales (NN) y support vector machine (SVM). Asimismo, se ha encontrado que los científicos tienden a realizar investigaciones con variables climáticas o demográficas para obtener mejores resultados. Se concluye que los métodos de aprendizaje automático que más se han utilizado son las redes neuronales de diferentes tipos: convolucional, recurrente, profunda y multicapa, y para la predicción de brotes de dengue predominaron los métodos de series de tiempo con LSTM y ARIMA, también se estableció que la tendencia es hacia la inclusión de variables climáticas y demográficas en los modelos de predicción.

Palabras-clave: Dengue, detección, machine learning, métodos de clasificación, algoritmos de clasificación, random forest, support vector machine, artificial neural networks

A systematic review of dengue classification models using machine learning

Abstract: Dengue is an arboviral disease that annually reports a large number of infected on the north coast and the Peruvian jungle. According to statistics, it is increasing yearly. This article aims to develop a systematic review of the scientific

literature on the study variables and the machine learning methods currently used for detecting dengue infection. The methodology used was PRISMA, initially mapping the literature of 274 scientific articles, leaving 33 articles selected for the systematic review. The results obtained are that the most used machine learning algorithms are neural networks (NN) and support vector machine (SVM). Likewise, it has been found that scientists tend to carry out research with climatic or demographic variables to obtain better results. It is concluded that the machine learning methods that have been used the most are neural networks of different types: convolutional, recurrent, deep, and multilayer, and for the prediction of dengue outbreaks the time series methods with LSTM and ARIMA were the predominant ones, it was also established that the trend is towards the inclusion of climatic and demographic variables in the prediction models.

Keywords: Dengue, detection, machine learning, classification methods, classification algorithms, random forest, support vector machine, artificial neural networks

1. Introducción

La inteligencia artificial ha brindado la generación de modelos y algoritmos que permiten predecir el resultado de una variable de estudio, de acuerdo a patrones encontrados en un conjunto de datos, en diferentes sectores de la actividad humana (Bravo et al., 2023; Cubas & Maquen-Niño, 2022; Maquen-Niño et al., 2022).

Recientemente, machine learning ha permitido mejorar la predicción de que un individuo tenga una enfermedad o padecimiento, en menor tiempo y con niveles cada vez más aceptables de acierto (Maquen-Niño et al., 2023). Es en este nuevo contexto, que cada vez son más los experimentos que se hacen en cuanto a detección de enfermedades arbovirósicas como es el dengue, ya que, al ser una enfermedad propagada por el mosquito *Aedes aegypti* y, en menor medida, *Aedes Albopictus*, de clima tropical, afecta a diferentes países en el mundo, y es declarada una enfermedad endémica en ciertas regiones que cíclicamente aparecen con fuertes rebrotes especialmente en épocas de lluvia intensa. Ecuador al igual que Perú por tener regiones de clima tropical sufren de rebrotes de dengue en ciertas épocas del año, así como en épocas del fenómeno del niño, lo que hace indispensable que haya un control epidemiológico de las regiones afectadas por el dengue (Pérez et al., 2022).

En este aspecto, la Organización mundial de la salud (OMS) expresa que la cantidad de contagios de dengue, o también conocido como Dengue Fever (DF) o Dengue hemorrhagic fever (DHF), en el mundo ha aumentado, y actualmente se estima que hay entre 100 y 400 millones de infecciones cada año, aunque solo el 80% de ellas son generalmente leves y asintomáticas, hay un porcentaje que pone la vida en riesgo (OMS, 2022). La inteligencia artificial y el internet de las cosas se han convertido en grandes aliados en el sistema de salud, como lo demuestra (Kamesh & Sivakumar, 2020), a través de una aplicación móvil que utiliza el algoritmo SVM, también para (Ho et al., 2020), utilizaron los datos médicos de pacientes confirmados con dengue y aplicaron modelos de predicción de la red neuronal profunda (DNN) de última generación, además árbol de decisión de convocatoria (DT) y los modelos de regresión logística (LR).

En otro trabajo de investigación de (Khan Dourjoy et al., 2021) se han utilizado support vector machine (SVM) y el algoritmo de clasificador de bosques aleatorios para la predicción del dengue, en cambio en (Siddiq et al., 2021), se desarrollaron y se comparó modelos lineales (regresión lineal) y tres modelos no lineales (regresión de vectores de soporte, regresión de árboles de decisión y regresión del bosque aleatorio) para determinar el modelo con la mayor precisión en la predicción de la transmisión de DF.

Cabe indicar que hay trabajos de investigación que se centran en los datos sintomatológicos de los casos de dengue, pero hay varios trabajos actualmente que se están centrando en los datos climatológicos como tenemos los trabajos de (Bokonda et al., 2021; McGough et al., 2021; Nguyen et al., 2022; Rahman et al., 2021; Xu et al., 2020; Yavari Nejad & Varathan, 2021), quienes han utilizado dichos datos para definir modelos de predicción mucho más precisos utilizando diversos algoritmos de machine learning. Se han aplicado estudios en diferentes lugares, como en Bangladesh donde el dengue ha aumentado, el trabajo de investigación de (Sarma et al., 2020), trabaja con los datos del paciente, que contiene información sobre el informe de diagnóstico del paciente, el historial médico y los síntomas, aplicaron algoritmos de aprendizaje automático, como: árbol de decisión (DT) y Random Forest (RF) en el modelo de clasificación propuesto, a esto se suma el trabajo de (Sukama et al., 2020), en esta investigación, aplican la red neuronal recurrente (RNN), para predecir incidentes de DHF en DKI Yakarta mediante el uso de datos históricos de casos DHF de 2009 a 2017. Además la investigación de (Sukama et al., 2020), quienes aplicaron la red neuronal recurrente (RNN), para predecir incidentes de DHF en DKI Yakarta mediante el uso de datos históricos de casos DHF de 2009 a 2017, otro caso fue el realizado por (Huang et al., 2020), quienes trabajaron con los datos de casos de dengue de un Hospital en Taiwan, entre julio a noviembre del 2015.

A nivel Latinoamérica, tenemos algunos trabajos en Brasil, como el de McGough et al. (2021), quienes tomaron información de los casos de dengue ocurridos en Brasil entre los años 2001 a 2007 y los datos climatológicos en los años 2000 a 2016 para predecir con mayor precisión esta enfermedad, se suma el trabajo de (Bauxell et al., 2021), quienes presentan un algoritmo para la predicción del riesgo de dengue en Brasil, se basan en técnicas de aprendizaje automático y utilizan variables ambientales y socioeconómicas para predecir el riesgo de propagación del dengue en todo Brasil. Otros estudios en Brasil son los realizados por (Aleixo et al., 2022; Koplewitz et al., 2022; Souza et al., 2022).

En este sentido, los hallazgos anteriores llevaron a la elaboración de las preguntas de investigación: ¿Cuáles son las variables de estudio en la detección del contagio por dengue?, ¿Cuáles son los modelos y algoritmos de machine learning usados? y ¿Cuáles son los rangos de precisión encontrados?

2. Métodos y materiales

La revisión sistemática fue desarrollada siguiendo la metodología PRISMA. Se utilizó las bases de datos indexadas: Scopus, Web of Science, IEEE Xplore y Google académico porque albergan artículos científicos que siguen criterios rigurosos aceptados por la comunidad científica, así como publicaciones de diferentes países que están a la

vanguardia en temas de machine learning orientados a la detección del dengue. Para realizar esta investigación, se contó como materia prima artículos científicos en inglés que fueron encontrados aplicando ecuaciones de búsqueda específica y sus sinónimos, con el uso de operadores booleanos AND y OR. Para el concepto: “Classification” se tomaron en cuenta los sinónimos: Classification models, prediction y para el concepto “Dengue” se tomaron en cuenta los sinónimos: Dengue fever, DENV. Por tanto, la ecuación de búsqueda quedó planteada en idioma inglés conforme se visualiza en la Tabla 1.

Idioma	Ecuación de búsqueda
English	(classification OR prediction OR “Classification models”) AND “machine learning” AND (dengue OR “dengue fever” OR DENV)

Tabla 1 – Ecuación de búsqueda

Se aplicó como filtro de búsqueda que los años de publicación de los artículos sean entre el año 2014 hasta el año 2022. De 274 publicaciones encontradas en ese rango de años, 184 pasaron la etapa de cribado, 96 cumplieron los criterios de elegibilidad y en la etapa final 33 cumplieron con los criterios de inclusión, como se observa en la Tabla 2. Los criterios de cribado corresponden a dos aspectos: la pertinencia y la accesibilidad. Por pertinencia se entiende que, en base a la lectura del resumen y conclusiones, los investigadores determinaron la pertinencia o no del artículo dentro de la presente investigación. Por accesibilidad se entiende que los archivos son de acceso abierto y descargables.

Los principales criterios de elegibilidad son los siguientes:

1. Los objetivos de la investigación están claramente especificados y se relacionan con la clasificación del dengue usando Machine Learning.
2. La investigación contiene aportes teóricos sobre las tendencias de la clasificación del dengue usando Machine Learning
3. La investigación define y explican claramente las variables de estudio para la detección del dengue.
4. Los modelos y algoritmos de clasificación de dengue obtenidos en la investigación son pertinentes para conseguir sus objetivos
5. El score obtenido en cada modelo de machine learning es mayor a 0.7

Database	Step 1: Identificación	Step 1: Cribado	Step 2: Elegibilidad	Step 3: Inclusión
Scopus	151	129	75	27
Web of science	76	35	10	1
IEEE Xplore	43	16	10	4
Google Scholar	4	4	1	1
Total	274	184	96	33

Tabla 2 – Bases de datos y fases desarrolladas para obtener información

La aplicación de criterios identificación, cribado, elegibilidad e inclusión puede verse en el diagrama de flujo PRISMA, detallado en la Figura 1.

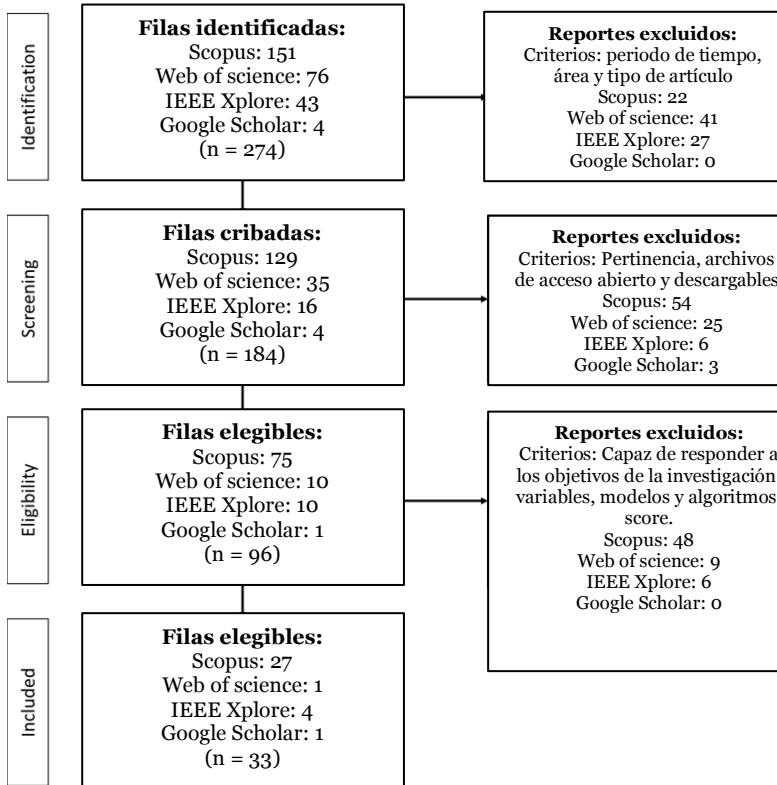


Figura 1 – Flujo de trabajo de PRISMA para la selección de los estudios relevantes

3. Resultados

Para realizar el análisis bibliométrico se utilizó la herramienta VOS-Viewer de los 33 artículos elegidos tomando como análisis las co-ocurrencias de todas las palabras claves.

Los resultados obtenidos del análisis bibliométrico realizado de la búsqueda inicial nos muestran tres cluster de las co-ocurrencias de todos los keywords de los artículos encontrados, como se observa en la Figura 2.

Del análisis de los cuatro clústeres encontrados, el clúster de color verde consta de 10 ítems y se centra en las técnicas para la predicción del dengue: machine learning, Decision Tree, Random Forest, time series analysis, prediction, forecasting. Otros conceptos relacionados que tienen gran nivel de ocurrencia son: dengue, aedes aegypti, virus infection, tropical countries.

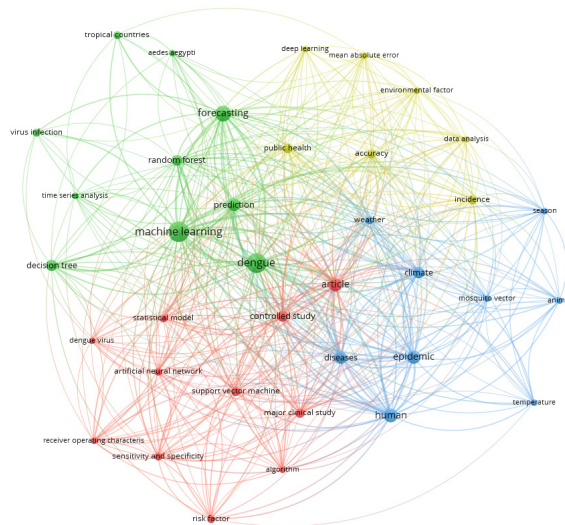


Figura 2 – Análisis de co-ocurrencias de palabras clave

El segundo clúster, es el clúster amarillo que consta de 7 ítems, centrando su atención en las medidas de evaluación del modelo como son: accuracy, mean absolute error, incidence, data analysis. Otros conceptos relacionados son: environmental factor, public health y deep learning.

El tercer clúster, consta de 11 ítems, es el clúster rojo, que hace referencia principalmente a modelos: statistical model, support vector machine, artificial neural networks, algorithm. Otros conceptos relacionados son: dengue virus, controled study, risk factor, receiver operating characterisc, major clinical study, article, sensitivity and especificity.

El cuarto clúster, es el clúster azul que consta de 9 ítems, que hace referencia a los atributos de las variables utilizadas para la predicción del dengue: weather, climate, season, temperature, epidemic, mosquito vector, diseases. También encontramos otros conceptos que tienen gran ocurrencia en los artículos estudiados como son: animal y human.

Un total de 33 artículos científicos fueron seleccionados para responder a las preguntas de investigación formuladas. Por cada artículo, se detalló la base de datos en la cual fue encontrado, el autor y el año de publicación, el país del cual la evidencia científica fue recolectada, sus palabras clave, así como las categorías o dimensiones investigadas. Los resultados fueron sistematizados en la Tabla 3, siendo las categorías denotadas por los siguientes números: Epidemiológica (Categoría 1), Clima (Categoría 2), Síntomas (Categoría 3), Demográficas (Categoría 4), Análisis de laboratorio Clínico (Categoría 5), socioeconómica (Categoría 6) y geográfica (Categoría 7).

DB	No.	Autor(es)	País de los datos	Título de artículo	Cate-goría
S C O P U S	1	(Appice et al., 2020)	Mexico	A Multi-Stage Machine Learning Approach to Predict Dengue Incidence: A Case Study in Mexico	1, 2
	2	(Srivastava et al., 2020)	India	An online learning approach for dengue fever classification	3
	3	(Huang et al., 2020)	Taiwan	Assessing the risk of dengue severity using demographic information and laboratory test results with machine learning	4, 5
	4	(Sukama et al., 2020)	DKI Jakarta	Comparing activation functions in predicting dengue hemorrhagic fever cases in DKI Jakarta using recurrent neural networks	3
	5	(Ho et al., 2020)	Taiwan	Comparing machine learning with case-control models to identify confirmed dengue cases	4, 5, 3
	6	(Sarma et al., 2020)	Bangladesh	Dengue Prediction using Machine Learning Algorithms	4, 5, 3
	7	(Xu et al., 2020)	China	Forecast of dengue cases in 20 Chinese cities based on the deep learning method	1, 2
	8	(Balamurugan et al., 2020)	India	Improved prediction of dengue outbreak using combinatorial feature selector and classifier based on the entropy-weighted score based optimal ranking	5
	9	(Raizada et al., 2020)	India	Vector-borne disease outbreak prediction by machine learning	1, 2
	10	(Benedum et al., 2020)	Iquitos, Puerto Rico y Singapore	Weekly dengue forecasts in Iquitos, Peru; San Juan Puerto Rico; and Singapore	1, 2
	11	(Khan Dourjoy et al., 2021)	Bangladesh – South Asia	A Comparative Study on Prediction of Dengue Fever Using Machine Learning Algorithm	3
	12	(McGough et al., 2021)	Brazil	A dynamic, ensemble learning approach to forecast dengue fever epidemic years in Brazil using weather and population susceptibility cycles	1, 2
	13	(Kaur et al., 2021)	India	A Machine Learning Approach for Predicting Dengue Outbreak	1, 2
	14	(Ferdousi et al., 2021)	Brazil	A Windowed Correlation-Based Feature Selection Method to Improve Time Series Prediction of Dengue Fever Cases	1, 2

DB	No.	Autor(es)	País de los datos	Título de artículo	Cate-goría
S C O P U S	15	(Yavari Nejad & Varathan, 2021)	Malaysia	Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction	1, 2
	16	(Bauxell et al., 2021)	Brazil	Machine learning techniques using environmental data from remote sensing applied to modeling dengue risk in Brazil	1, 6, 2
	17	(Rahman et al., 2021)	Thailand	Mapping the spatial distribution of the dengue vector <i>Aedes aegypti</i> and predicting its abundance in northeastern Thailand using machine-learning approach	6, 2, 3
	18	(Choubey et al., 2021)	San Jose e Iquitos	Soft computing techniques for dengue prediction	2
	19	(Tanawi et al., 2021)	DKI Jakarta	Support Vector Regression for Predicting the Number of Dengue Incidents in DKI Jakarta	2
	20	(Nguyen et al., 2022)	Vietnam	Deep learning models for forecasting dengue fever based on 2 data in Vietnam	1
	21	(Abdualgalil et al., 2022)	Yemen	Early Diagnosis for Dengue Disease Prediction Using Efficient Machine Learning Techniques Based on Clinical Data	4, 3, 5
	22	(Ochida et al., 2022)	New Caledonia	Modeling present and future climate risks of dengue outbreak, a case study in New Caledonia	2
	23	(Koplewitz et al., 2022)	Brazil	Predicting dengue incidence leveraging internet-based data sources. A case study in 20 cities in Brazil	1, 7, 2
	24	(Souza et al., 2022)	Brazil	Predicting dengue outbreaks in Brazil with manifold learning on 2 data	2
	25	(Aleixo et al., 2022)	Brazil	Predicting Dengue Outbreaks with Explainable Machine Learning	1, 4, 2
	26	(Kumar et al., 2022)	Bangladesh	Prediction of dengue incidents using hospitalized patients, meteorological and socioeconomic data in Bangladesh: A machine learning approach	1, 2
	27	(Nalini et al., 2022)	San Juan e Iquitos	Prediction of dengue infection using Machine Learning	2

DB	No.	Autor(es)	País de los datos	Título de artículo	Cate-goría
WoS	28	(Salim et al., 2021)	Malaysia	Prediction of dengue outbreak in Selangor Malaysia using machine learning techniques	1, 2
	29	(Baker et al., 2021)	San Juan e Iquitos	Forecasting Dengue Fever Using Machine Learning Regression Techniques	3, 2
IEEEExplore	30	(Singh, 2021)	India	Particle Swarm Optimization assisted Support Vector Machine based Diagnostic System for Dengue prediction at the early stage	3
	31	(Siddiq et al., 2021)	Saudi Arabia	Predicting Dengue Fever Transmission Using Machine Learning Methods	1, 2
	32	(Rana et al., 2022)	Kaggle (internet repository)	Dengue Fever Prediction using Machine Learning Analytics	3
Google	33	(Chattopadhyay & Chattopadhyay, 2022)	India	VIRDOC: A VIRTUAL DOCTOR to predict dengue fatality.	3

Tabla 3 – Base de datos, lista de autores, país, título y categorías exploradas por artículo.

Del total de 33 artículos seleccionados, algunos fueron de Brasil (6), India (6), Bangladesh (3), Malaysia (2), Taiwan (2), DKI Jakarta (2), San Juan e Iquitos (2), San José e Iquitos (1), Iquitos, Puerto Rico y Singapore (1), Arabia Saudita (1), Vietnam (1), Thailandia (1), Yemen (1), New Caledonia (1), México (1), China (1) y Kaggle (1). Por otro lado, en las categorías exploradas por los autores, tenemos clima (21), epidemiológica (15), síntomas (11), análisis de laboratorio clínico (5), demográficas (5), socioeconómica (2) y geográfica (1) correspondientes a las variables de estudio, conforme se visualiza en la Figura 3.

Un análisis más profundo de cada Artículo científico permitió identificar aspectos relacionados a las variables/características seleccionadas, fuente de datos, algoritmo / método de machine learning y, método de validación en cada artículo (ver Tabla 4).

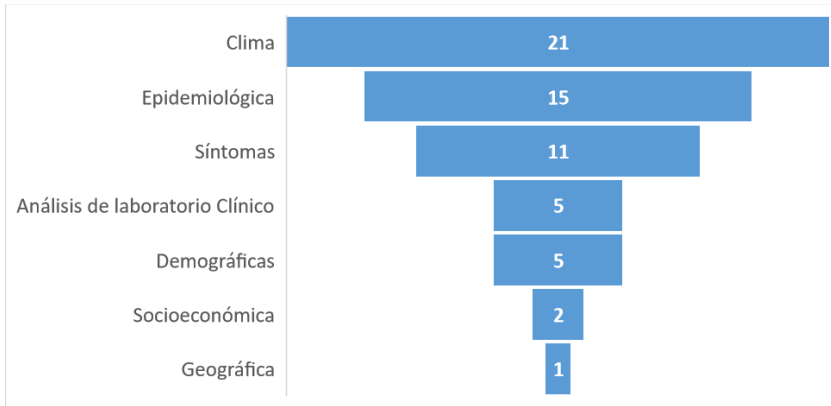


Figura 3 – Número de artículos por categoría

4. Discusión de resultados

La revisión sistemática permitió seleccionar un total de 33 artículos, muchos de los cuales son de Brasil, India, Bangladesh e Iquitos. La categoría más estudiada por los autores es la de Clima (21), seguido por la categoría epidemiológica (15) y la categoría síntomas (11), demostrando que existe una gran tendencia a los estudios con variables de factores climáticos que ayudan a mejorar la predicción de brotes de dengue y la detección de la enfermedad de acuerdo con la ubicación geográfica del paciente.

Esto debido a que el vector mosquito que produce la enfermedad, es de clima tropical y subtropical, lo cual está asociado con variables climatológicas como la temperatura ambiental, la humedad y sobre todo el nivel de precipitación. De esa forma, si es que hay temporadas de lluvias, el agua muchas veces se queda empozada y es criadero de zancudos propagadores del dengue.

Otra de las categorías que también es predominante en los estudios y ocupa el segundo lugar, es la de factores epidemiológicos, que está ligado altamente al control sanitario de cada país en las provincias donde el dengue es una enfermedad endémica, que tiene que ver con el registro de contagiados, hospitalizados y muertos que reportan cada cierto tiempo los hospitales y centros de salud, y que para ayudar a los modelos de machine learning deberían tener la misma distribución temporal, sea diaria, semanal o mensual, pero que no haya periodos en los que no hubo reporte de datos, ya que esto afecta la continuidad de los datos y para reemplazar esos faltantes se tienen que realizar estimaciones a través de métodos predictivos.

La categoría Síntomas que es la que ocupa el tercer lugar, se caracteriza por tener investigaciones que han utilizado datos descargados de repositorios oficiales de un determinado país y en su minoría base de datos construidas por los investigadores de datos obtenidos de clínicas y hospitales. Entre los síntomas que tienen mayor incidencia

Nº	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
1	Dos variables: Número de casos de dengue y dengue hemorrágico. Temperatura media del aire (en °C) a nivel mensual por cada uno de los 32 estados	1, 2	Serie de Tiempo de ambas variables recolectados durante el periodo 1985-2010 en 32 estados federales de México. Training (1985-2009) y test (2010)	Time Series Forecasting: AutoRegressive Integrated Moving Average (auto.arima) y Vector Autoregression (VAR). Regression Methods (M5), Support Vector Regression (SVR), k-Nearest Neighbor (kNN) y la estrategia de aprendizaje propuesta de múltiples etapas llamada AutoTIC-NN. AutoTIC-NN demostró el menor RMSE en la mayor cantidad de estados evaluados.	Root Mean Square Error (RMSE).
2	14 variables correspondientes a Síntomas de pacientes e investigaciones de laboratorio de los pacientes	3	182 muestras de pacientes, 148 positivos y 34 negativos. (training 80% y test 20%)	Support Vector Machine (SVM), Random Forest (RF), así como múltiples clasificadores de Weka: Logistic, SGD, SGD text, Simple Logistic, SMO, Voted perceptron, Adaboost, Random subspace, Bayes Net, Naive Bayes (NB) y LWL. Clasificadores en Línea: aROW, PA2, ALMA, a JELLIP, CW, POE, aROMA, ROMA. (RF obtuvo 82.41%)	F - MEASURE
3	Variables demográficas y análisis de laboratorio clínicos	4, 5	798 casos (training set 90%, test set 10%)	Logistic Regression, SVM, RF, gradient boosting machine (GBM), Artificial Neural Network (ANN) (ANN: ROC: 0.8324 ± 0.0268, Accuracy: (0.7523 ± 0.0273))	10 - fold cross validation approach, ROC and Accuracy
4	Incidencias de Dengue hemorrágico	3	Data de Dengue hemorrágico (DHF) de 418 semanas de 5 municipalidades de Jakarta entre el 2009 y 2017 (Training 70% y test 30%)	Recurrent Neural Network (RNN) y Backpropagation Through Time (BPTT). Los resultados de la predicción de DHF en las cinco municipalidades de Jakarta obtuvieron mejores resultados en el RMSE y RME con la función de activación Sigmoid.	RMSE and MSE (The Mean Squared Error).
5	18 variables, siendo: demográficas, clínicas (signos vitales y comorbilidades), de análisis de laboratorio.	4, 5, 3	4894 casos de atención por emergencia en un Hospital	Logistic Regression, Deep Neural Network, Decision Tree (DT) (DNN fue el mejor con 86.95%±0.45%)	10-fold cross validation process

N°	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
6	23 atributos: demográficos, síntomas y análisis de laboratorio	4, 5, 3	101 000 casos de Bangladesh de dos ciudades principales densamente pobladas (Dhaka y Chittagong)	Sistema clasificador binario (0,84)	ROC curve
7	Número de casos de dengue humano por mes (variable epidemiológica) y 9 variables meteorológicas mensuales.	1, 2	Datos meteorológicos y Dato epidemiológico mensual del año 2005 al 2018 de 20 ciudades de China (Training data 2005- al 2016, Testing Data 2017-2018)	Redes neuronales recurrentes de memoria a corto plazo (LSTM) y Redes neuronales recurrentes de memoria a corto plazo con aprendizaje de transferencia (LSTM - TL) y su comparación con otros modelos: BPNN; Backpagation Neural Network; GAM; Generalized Additive Model; SVR; Support Vector Regression; GBM; Gradient Boosting Machine (LSTM 0.9087 / 1.1600)	RMSE (The Root Mean Squared Error), RRSE (The Root Relative Squared Error)
8	20 variables correspondientes a los atributos de un análisis de laboratorio de sangre	5	480 muestras de pacientes con dengue	EWSORA Algoritmo de clasificación óptima basado en puntaje ponderado de entropía. Genetic Algorithm (GA), Particle Swarm Algorithm (PSO), Correlation based Feature Selector (CFS) Modelos de ML: J48, NB, SVM y Multilayer perceptron (MLP), regresión y series de tiempo (EWSORA con MLP con 98,7%)	cross validation
9	Casos positivos del 2013 al 2017 para Chikungunya, malaria y dengue. También Variables meteorológicas: temperatura, humedad y precipitación.	1, 2	541-5958 datos de casos positivos y datos meteorológicos de 36 estados del 2013 al 2017 fusionados. Tres posibles salidas: riesgo alto, moderado o bajo. (Training: 60%y Test 40%).	Convolutional Neural Network multimodal disease outbreak prediction CNN- MDOP y Artificial Neural Network multimodal disease outbreak prediction ANN - MDOP (CNN accuracy: 0.88)	Prediction accuracy
10	Recuentos de casos y brotes semanales de dengue en Iquitos, Perú; San Juan, Puerto Rico; y Singapur de 1990 a 2016. Variables meteorológicas de esos tres lugares (16 variables)	1, 2	Se crearon 500 conjuntos de entrenamiento que se usó para entrenar cada modelo y se promediaron las predicciones	Algoritmo ML predictivo: RF, RF con Algoritmo de señalización univariante (RF-UFA) Regresión de Poisson y regresión logística. Modelo de promedio móvil integrado autorregresivo (ARIMA) y modelo estacional ARIMA (SARIMA) (Modelos de ML - RF tenían un 21 % y un 33 % menos de error)	Cross Validation with fold =10 MAE

Nº	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
11	15 variables de sintomatología de Dengue, buscándose predicción para el próximo año	3	Encuestas online a más de 600 pacientes de Bangladesh y el método de topografía. Se han recopilado casi 1047 datos. Training 80 % y test 20%	SVM y el algoritmo clasificador de Random Forest (RFC), (SVM 0.695)	cross validation
12	Variables de pacientes con dengue y variables climáticas como: temperatura y precipitación diaria	1, 2	Datos sobre casos anuales de dengue (Ministerio de Salud de Brasil) para 2001–2017 y sobre temperatura y precipitación diaria (GMAO-NASA) para 2000–2016, para 20 municipios endémicos de dengue en Brasil	Contempla 11 modelos de aprendizaje ensablado que usan SVM. En función de si es un año epidémico o no, el 'strength' de su modelo es del 63% o el 71% respectivamente.	Accuracy, Time-series
13	Variables meteorológicas y el conteo de casos de Dengue como variable epidemiológica.	1, 2	Datos obtenidos del Departamento Meteorológico de la India, Pune entre 2006-2015.	Hidden Markov Model (HMM) para datos de series de tiempo. La precisión obtenida para los cuatro distritos es 97% para Bathinda, 92% para Patiala, 97% para Ludhiana y 86% para Amritsar.	Accuracy
14	Casos de dengue y variables climáticas	1, 2	78 municipios en Brasil con casos 71348, 58424, 45319, 36743, 27103 Training 50%, validation 30% y test 20%	Recurrent neural networks (RNN), LSTM, GRU y el modelo Lineal. (GRU: total de casos MAE 0.13 y para prueba de varias ubicaciones MAE rango de 0.2435 a 0.2569)	Mean absolute error (MAE)
15	Variables sobre casos de Dengue (incidencias y casos confirmados) y variables climáticas como son Temperatura, humedad relativa y cantidad de lluvia.	1, 2	Datos obtenidos de dos fuentes oficiales. El incidente de dengue y los casos confirmados semanales del portal del Ministerio de Salud de Malasia (MOH) y, los datos climáticos se obtienen del Departamento Meteorológico de Malasia (MMD) en el periodo 2010 al 2013	Bayes Network models (BN), SVM, Radial Basis Function (RBF Tree), decision table (DT) y NB. (Bayes Networks con 92-35% de accuracy)	cross validation (ten times) with TempeRain factor (TRF)

N°	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
16	Variables clínicas de pacientes con dengue en Brasil. Variables ambientales y socioeconómicas	1, 6, 2	Datos de 2010 a 2013 de la distribución mensual de episodios de dengue del Sistema de Información de Enfermedades Notificables (SINAN), desarrollado por el Ministerio de Salud de Brasil. Datos ambientales de las misiones satelitales de observación de la Tierra (Training 80%, test 20%)	DT, RF (R2: 0.7635 y accuracy: 89,49% usando Index B) y neural networks (R2: 0.5041 and accuracy: 92,40% usando index B y un accuracy: 79.21% usando el index A). RF es el de mejor resultado.	R-squared (R2) and accuracy
17	Variables socioeconómicas, variables climáticas y síntomas de dengue	6, 2, 3	Tailandia de enero a diciembre del 2019	Modelos de aprendizaje supervisados, regresión logística (LR), SVM, vecino K-Nearest (KNN), red neuronal artificial (ANN) y RF. RF, que usó factores socioeconómicos, KAP y de paisaje como entradas, tuvo la mayor precisión del modelo de predicción, seguida por ANN, kNN, LR y SVM (socioeconómico: CA = 0,86 y F1 = 0,85; KAP: CA = 0,92 y F1 = 0,90 y paisaje: CA = 0,89 y F1 = 0,87)	AUC (área bajo la curva ROC), CA (precisión de clasificación) y puntaje F1
18	Variables climáticas cualitativas y cuantitativas	2	Datos de DengAI Data set y los datos climáticos de dos ciudades San José e Iquitos	XGBoost Regressor Model, Negative Binomial Regressor Model, DT Regressor Model, ARIMA Model. En base a los valores MAE y las compensaciones, el regresor binomial negativo (NBR) es el modelo con mejor rendimiento.	MAE, RMSE
19	Las variables meteorológicas, incluidos las variables sobre la lluvia, la humedad y la temperatura	2	Datos de Cinco regiones en DKI Yakarta desde el 6 de enero de 2009 hasta el 25 de septiembre de 2017, proporcionados por la Agencia de Meteorología, Climatología y Agencia Geofísica (BMKG)	Support Vector Regression (Cross-correlation, Augmented Dickey-Fuller test para calcular el lapso entre dos series de tiempo estacionarias). SVM con núcleo lineal dio errores relativamente pequeños en la predicción de incidentes de dengue (MAE < 8 y RMSE <10).	MAE, RMSE

Nº	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
20	12 factores meteorológicos en el mismo período, incluidas medidas de temperatura, lluvia, humedad, evaporación y horas de sol, casos y muertes por tasa mensual del dengue	1	Casos de dengue de 20 provincias/ciudades pertenecientes a tres regiones principales de Vietnam entre 1997 y 2016 del Instituto Nacional de Higiene y Epidemiología de Vietnam. Datos meteorológicos del Instituto de Meteorología, Hidrología y Cambio Climático de Vietnam. Training (1997 a 2013) y test (2014 a 2016) para provincia	Long short-term memory neural network based on the attention mechanism (LSTM-ATT), Long short-term memory neural network LSTM, Convolutional Neural Network (CNN), Transformer (TF) (LSTM-ATT, RMSE: 1,60 y MAE: 1,90)	MAE, RMSE
21	21 atributos: edad, sexo, fiebre, síntomas y variables clínicas.	4, 3, 5	Se obtuvo del Centro de Monitoreo Epidemiológico (EMC) - Oficina de Salud Pública y Población (PHPO) - Taiz City, Yemen del 2017 al 2019.	Técnicas eficientes de aprendizaje automático (EMLT): K-Nearest Neighbor (KNN), clasificador de impulso de gradiente (GBC), clasificador de árbol adicional (ETC), impulso de gradiente extremo (XGB) y máquina de impulso de gradiente de luz (LightGBM). El mejor modelo fue ETC con una accuracy del 99,03%, f1-score del 99,04%, precisión del 98,92% y recall del 99,17%.	Accuracy, f1-score, recall, precision, AUC, and runtime
22	La lluvia diaria (RR) y la temperatura máxima (TX)	2	Los datos demográficos de Nueva Caledonia de los censos de 1969, 1976, 1983, 1989, 1996, 2004, 2009, 2014 and 2019. Los datos climáticos de una estación meteorológica de Météo-France de 1970 a 2020 en Faubourg Blanchot, Nouméa.	SVM. Se presenta un método para evaluar la probabilidad semanal de un brote de dengue para un lugar específico y proyectar este riesgo en un escenario de cambio climático.	Leave Time Out - Cross Validation (LTO - CV)

N°	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
23	Variables epidemiológicas, geográficas y climáticas	1, 7, 2	Incidencias semanales de años anteriores, variables climáticas semanales en Brasil y datos de búsqueda en Internet en tiempo real. Periodo de enero del 2010 a Julio del 2016.	RF y L1 – based (LASSO Regression). Se determinó que RF es más consistente y robusto que LASSO.	RMSE, relative RMSE, R2
24	Variables climáticas	2	Seis ciudades de Brasil	SVM. Random Guess (R-Guess) accuracy 0.8	Test Mean Accuracy
25	Variables epidemiológicas, meteorológicas, sociodemográficas	1, 4, 2	Datos del Sistema Nacional de Información sobre Enfermedades Notificables (SINAN), datos meteorológicos y datos sociodemográficos del Instituto Brasileño de Geografía y Estadísticas	Gradient boosting decision tree algorithm (CatBoost), accuracy 90%	MAE, RMSE, and R2
26	Casos de dengue del 2012 al 2019 y Variables climatológicos de 11 ciudades de Bangladesh	1, 2	11 distritos diferentes de Bangladesh	Modelo de aprendizaje automático, regresión lineal múltiple (MLR) y regresión de vectores de soporte (SVR). Support Vector Regression (SVR) 75% de precisión.	Precision, mean absolute error (MAE) of models
27	Temperatura máxima y mínima, temperatura promedio, precipitación total de la estación meteorológica climática GHCN de la NOAA.	2	Datos de las ciudades de San Juan e Iquitos	Regresión lineal, Series de tiempo, Heat Mapping, Moving mean. Obtuvo mejores resultados es Linear Regression con un MAE de 20.77	Mean absolute error (MAE)
28	Variables Epidemiológicas y Variables climáticas (temperatura, velocidad del viento, humedad y lluvia).	1, 2	Los datos contienen 5 años (2013 a 2017) de casos semanales para cinco distritos en Selangor, Malasia: Gombak, Hulu Selangor, Hulu Langat, Klang y Petaling.	Classification and Regression Tree (CART), Artificial Neural Network (MLP), SVM (LINEAR, POLYNOMIAL, RBF), and Bayes Network (TAN). SVM (núcleo lineal) exhibió el mejor rendimiento de predicción con una exactitud de 70 %.	Accuracy, sensitivity, and precision

Nº	Variables/ características	Categoría	Fuente de datos	Algoritmo/ método	Método de validación
29	Variables sintomatológicas y climatológicas (Location, Temperature, Precipitation, Humidity, and Vegetation)	3, 2	DrivenData Website de dos ciudades San Juan e Iquitos	Machine Learning Regression, regression neural network, RF, Negative binomial regression, Poisson Regression, Decision Tree Regression (DTR), NB, SVM, k-nearest neighbor (k-NN) Regression, AdaBoost, Bagging with decision tree. Obtuvo mejores resultados Recurrent Neural Network (RNN) con memoria a largo plazo (LSTM) con una precisión del 94%	Mean absolute error (MAE)
30	Variables sintomatológicas	3	Conjuntos de datos de clasificación de dengue estándar	J48, RF, DT y Algoritmo propuesto (PSO - SVM). Obtuvo mejores resultados el algoritmo propuesto con 87% de accuracy	MAE
31	Variables epidemiológicas y variables climatológicas: Temperatura y humedad	1, 2	Casos de DF reportados para Jeddah City, Arabia Saudita. La temperatura y humedad del Research Data Archive at the National Center for Atmospheric Research	Regresión lineal y tres modelos no lineales (regresión de vectores de soporte, regresión de árboles de decisión y regresión del bosque aleatorio). Obtuvo mejor resultados Support Vector Classification con 76 % de accuracy	Precision
32	Variables sintomatológicas	3	Conjunto de datos con 124 muestras y 8 características descargado de Kaggle	Análisis de aprendizaje automático (DFES-MLA) utiliza clasificadores de árbol de decisión (DT) y bosque aleatorio (RF). Obtuvo mejor resultado Random Forest con un accuracy de 0.97.	Accuracy.
33	Variables sintomatológicas	3	Conjuntos de datos de clasificación de dengue estándar	Regresión lineal múltiple (MLR), ANOVA. Obtuvo mejores resultados Multiple Linear Regression con 75% de accuracy.	accuracy and performance

Tabla 4 – Variables/características seleccionadas, categoría, fuente de datos, algoritmo / método de machine learning y método de validación en cada artículo.

en las investigaciones podemos mencionar: Fiebre, dolor de cabeza, dolor corporal, artralgias, mialgias, náuseas, mareo, hemorragia, entre otros.

A nivel mundial existe un consenso para clasificar los tipos de contagios por dengue, siendo tres: Dengue sin signos de alerta, Dengue con signos alerta y Dengue hemorrágico o grave, y siendo la última la más peligrosa para la salud, y aunque es el que se da en muy baja proporción, puede haber riesgo de muerte de no tratarse a tiempo, de las 33 investigaciones seleccionadas, 2 investigaciones han abordado detalladamente la detección del dengue hemorrágico.

Con respecto a las técnicas y algoritmos de machine learning se puede denotar que hay una fuerte tendencia al uso de redes neuronales (17 estudios), ya que, debido a su variedad de topologías, son muy adaptables en diferentes tipos de investigación tanto para clasificación, como para regresión. Dentro de los modelos de redes neuronales utilizados podemos citar: ANN (7 estudios), LSTM (3 estudios), MLP (2 estudios), CNN (2 estudios), RNN (1 estudio), BPNN (1 estudio) y Perceptrón (1 estudio). Otras de las técnicas de machine learning que tuvieron gran incidencia fueron: SVM (12 estudios) seguida por RF (8 estudios), DT (6 estudios), Naive Bayes (4 estudios) para la detección del contagio de la enfermedad. Otras investigaciones se enfocaron hacia la regresión de datos, siendo las de mayor incidencia Logistic regression (5 estudios), Support vector regression (5 estudios), Linear regression (4 estudios) y Machine learning regression (4 estudios). También se utilizó métodos de series de tiempo con el uso de LSTM (3 estudios) y ARIMA y sus variantes (4 estudios).

Considerando la precisión de los modelos para la detección del dengue, se encontró que 9 artículos tuvieron como accuracy un valor mayor o igual que 70% y menor de 80%, 6 artículos tuvieron un valor mayor o igual a 80% y menor de 90%, 11 artículos tuvieron un valor mayor o igual a 90% y menor o igual a 100%, presentándose también que 7 artículos no especificaban el valor de precisión.

5. Conclusiones

Se llega a la conclusión que los métodos de machine learning que son más utilizados en la detección del contagio de dengue son las redes neuronales siendo las más ampliamente utilizadas las redes neuronales artificiales convolucionales y perceptrón multicapa. Para lo que corresponde a predicción de brotes de dengue se utilizaron preferentemente los modelos de serie de tiempo con LSTM y los modelos ARIMA con sus variantes, estos últimos más ligados a modelos estadísticos.

Existe una fuerte tendencia hacia la inclusión de variables climáticas en los modelos de machine learning que ayuden a mejorar la predicción, debido a que el mosquito que lo transmite tiene un hábitat con variables climáticas establecidas. Así mismo, la segunda categoría con mayor incidencia fue la epidemiológica debido a que los países están dando mayor importancia al control y monitoreo de enfermedades arbovirósicas como lo es el dengue, a través de la publicación de tasas relacionadas con las infecciones en los portales web oficiales, siendo esta data de fácil acceso a los investigadores.

Por último, se debe señalar que una gran cantidad de artículos (11) mostraron un rango de precisión mayor o igual al 90%, lo que se puede considerar como un rango

de precisión aceptable para modelos de machine learning que busquen detectar la enfermedad del dengue.

Referencias

- Abdualgalil, B., Abraham, S., & Ismael, W. M. (2022). Early Diagnosis for Dengue Disease Prediction Using Efficient Machine Learning Techniques Based on Clinical Data. *Journal of Robotics and Control*, 3(3), 257–268. <https://doi.org/10.18196/jrc.v3i3.14387>
- Aleixo, R., Kon, F., Rocha, R., Camargo, M. S., & Camargo, R. Y. (2022). Predicting Dengue Outbreaks with Explainable Machine Learning. In *2022 22nd International Symposium on Cluster, Cloud and Internet Computing (CCGrid) 2022*, (940–947). <https://doi.org/10.1109/CCGrid54584.2022.00114>
- Appice, A., Gel, Y. R., Iliev, I., Lyubchich, V., & Malerba, D. (2020). A Multi-Stage Machine Learning Approach to Predict Dengue Incidence: A Case Study in Mexico. *IEEE Access*, 8, 52713–52725. <https://doi.org/10.1109/ACCESS.2020.2980634>
- Baker, Q. B., Faraj, D., & Alguzo, A. (2021). Forecasting Dengue Fever Using Machine Learning Regression Techniques. In *2021 12th International Conference on Information and Communication Systems, ICICS 2021*, May, 157–163. <https://doi.org/10.1109/ICICS52457.2021.9464619>
- Balamurugan, S. A., Mallick, M. S. M., & Chinthana, G. (2020). Improved prediction of dengue outbreak using combinatorial feature selector and classifier based on entropy weighted score based optimal ranking. *Informatics in Medicine Unlocked*, 20(March), 100400. <https://doi.org/10.1016/j.imu.2020.100400>
- Bauxell, J., Vall-Ilossera, M., & Gurgel, H. (2021). Machine learning techniques using environmental data from remote sensing applied to modeling dengue risk in Brazil. In *International Geoscience and Remote Sensing Symposium*, (4608–4611). <https://doi.org/10.1109/IGARSS47720.2021.9554704>
- Benedum, C. M., Shea, K. M., Jenkins, H. E., Kim, L. Y., & Markuzon, N. (2020). Weekly dengue forecasts in Iquitos, Peru; San Juan, Puerto Rico; and Singapore. *PLoS Neglected Tropical Diseases*, 14(10), 1–26. <https://doi.org/10.1371/journal.pntd.0008710>
- Bokonda, P. L., Ouazzani-Touhami, K., & Souissi, N. (2021). Which Machine Learning method for outbreaks predictions? In *2021 IEEE 11th Annual Computing and Communication Workshop and Conference, CCWC 2021*, ML, (825–828). <https://doi.org/10.1109/CCWC51732.2021.9376061>
- Bravo, J., Alarcón, R., Valdivia, C., & Serquén, O. (2023). Application of Machine Learning Techniques to Predict Visitors to the Tourist Attractions of the Moche Route in Peru. *Sustainability*, 15(11), 8967. <https://doi.org/10.3390/SU15118967>
- Chattopadhyay, A. K., & Chattopadhyay, S. (2022). VIRDOCD: A VIRtual DOCTOR to predict dengue fatality. *Expert Systems*, 39(1), 1–17. <https://doi.org/10.1111/exsy.12796>

- Choubey, D. K., Mishra, A., Pradhan, S. K., & Anand, N. (2021). Soft computing techniques for dengue prediction. In *Proceedings - 2021 IEEE 10th International Conference on Communication Systems and Network Technologies, CSNT 2021*, (648–653). <https://doi.org/10.1109/CSNT51715.2021.9509706>
- Cubas, J. V., & Maquen-Niño, G. L. E. (2022). Machine learning model in the detection of phishing websites. *RISTI - Revista Ibérica de Sistemas e Tecnologias de Informação*, (E52), 161–173. <https://www.scopus.com/record/display.uri?eid=2-s2.0-85143599912&origin=resultslist&sort=plf-f>
- Ferdousi, T., Cohnstaedt, L. W., & Scoglio, C. M. (2021). A Windowed Correlation-Based Feature Selection Method to Improve Time Series Prediction of Dengue Fever Cases. *IEEE Access*, 9, 141210–141222. <https://doi.org/10.1109/ACCESS.2021.3120309>
- Ho, T. S., Weng, T. C., Wang, J., Der Han, H. C., Cheng, H. C., Yang, C. C., Yu, C. H., Liu, Y. J., Hu, C. H., Huang, C. Y., Chen, M. H., King, C. C., Oyang, Y. J., & Liu, C. C. (2020). Comparing machine learning with case-control models to identify confirmed dengue cases. *PLoS Neglected Tropical Diseases*, 14(11), 1–21. <https://doi.org/10.1371/journal.pntd.0008843>
- Huang, S. W., Tsai, H. P., Hung, S. J., Ko, W. C., & Wang, J. R. (2020). Assessing the risk of dengue severity using demographic information and laboratory test results with machine learning. *PLoS Neglected Tropical Diseases*, 14(12) 1–19. <https://doi.org/10.1371/journal.pntd.0008960>
- Kamesh, R., & Sivakumar, N. (2020). Design and implementation of realtime detection of dengue using machine learning. *Journal of Advanced Research in Dynamical and Control Systems*, 12(5 Special Issue), 1013–1025. <https://doi.org/10.5373/JARDCS/V12SP5/20201851>
- Kaur, L., Bhatia, R., & Nagpal, S. (2021). A Machine Learning Approach for Predicting Dengue Outbreak. In *Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021*, (60–65). <https://doi.org/10.1109/ICAC3N53548.2021.9725539>
- Khan Dourjoy, S., Rabbani Rafi, A., Nasrin Tumpa, Z., & Saifuzzaman, M. (2021). A Comparative Study on Prediction of Dengue Fever Using Machine Learning Algorithm. In *Lecture Notes in Networks and Systems* (Vol. 127). https://doi.org/10.1007/978-981-15-4218-3_15
- Koplewitz, G., Lu, F., Clemente, L., Buckee, C., & Santillana, M. (2022). Predicting dengue incidence leveraging internet-based data sources. A case study in 20 cities in Brazil. *PLoS Neglected Tropical Diseases*, 16(1). <https://doi.org/10.1371/JOURNAL.PNTD.0010071>
- Kumar, S., Rahman, M., Howlader, A., Siddiqi, U., Mohi, K., Borhan, R., & Rahman, E. (2022). Prediction of dengue incidents using hospitalized patients, metrological and socio-economic data in Bangladesh_ A machine learning. *Plos One*. <https://doi.org/https://doi.org/10.1371/journal.pone.0270933>

- Maquen-Niño, G. L. E., Sandoval-Juarez, A. A., Veliz La Rosa, R. A., Carrión-Barco, G., Adrianzén-Olano, I., Vega-Huerta, H., & De-La-Cruz-VdV, P. (2023). Brain Tumor Classification Deep Learning Model Using Neural Networks. *International Journal of Online and Biomedical Engineering*, 19(9), 81–92. <https://doi.org/10.3991/IJOE.V19I09.38819>
- Maquen-Niño, G. L. E., Villegas-Cubas, J. E., Celis-Bravo, P. J., Carrión-Barco, G., & Adrianzen-Olano, I. (2022). Integrative methodology model for the creation of pathfinding systems. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, (E51), 46–57. https://www.researchgate.net/publication/370442097_Integrative_methodology_model_for_the_creation_of_pathfinding_systems
- McGough, S., Clemente, L., Kutz, N., & Santillana, M. (2021). A dynamyc, ensemble learning approach to forecast dengue fever epidemic years in Brazil using weather and population susceptibility cycles. *RSIF*. 18(179). <https://doi.org/10.1098/rsif.2020.1006>
- Nalini, C., Shanthakumari, R., Prasanna, R. V., Nikilesh, A., & Pranav, S. M. N. (2022). Prediction of dengue infection using Machine Learning. In 2022 International Conference on Computer Communication and Informatics, ICCCI 2022; <https://doi.org/10.1109/ICCCI54379.2022.9740924>
- Nguyen, V.-H., Tuyet-Hanh, T. T., Mulhall, J., Minh, H. V., Duong, T. Q., Chien, N. V., Nhung, N. T. T., Lan, V. H., Minh, H. B., Cuong, D., Bich, N. N., Quyen, N. H., Linh, T. N. Q., Tho, N. T., Nghia, N. D., Anh, L. V. Q., Phan, D. T. M., Hung, N. Q. V., & Son, M. T. (2022). Deep learning models for forecasting dengue fever based on climate data in Vietnam. *PLOS Neglected Tropical Diseases*, 16(6), e0010509. <https://doi.org/10.1371/journal.pntd.0010509>
- Ochida, N., Mangeas, M., Dupont-Rouzeyrol, M., Dutheil, C., Forfait, C., Peltier, A., Descloux, E., & Menkes, C. (2022). Modeling present and future climate risk of dengue outbreak, a case study in New Caledonia. *Environmental Health: A Global Access Science Source*, 21(1), 1–10. <https://doi.org/10.1186/s12940-022-00829-z>
- OMS. (2022). Dengue y dengue grave. <https://www.who.int/es/news-room/fact-sheets/detail/dengue-and-severe-dengue>
- Pérez, L. G. A., García, M. I. B., Babici, V. R., & Monge, E. J. A. (2022). Dengue fever after the covid-19 pandemic: a statistical software analysis of its current situation in Ecuador. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, (E53), 362–374.
- Rahman, M. S., Pientong, C., Zafar, S., Ekalaksananan, T., Paul, R. E., Haque, U., Rocklöv, J., & Overgaard, H. J. (2021). Mapping the spatial distribution of the dengue vector *Aedes aegypti* and predicting its abundance in northeastern Thailand using machine-learning approach. *One Health*, 13(December). <https://doi.org/10.1016/j.onehlt.2021.100358>

- Raizada, S., Mala, S., & Shankar, A. (2020). Vector borne disease outbreak prediction by machine learning. In Proceedings of the International Conference on Smart Technologies in Computing, Electrical and Electronics, ICSTCEE 2020, (213–218). <https://doi.org/10.1109/ICSTCEE49637.2020.9277286>
- Rana, S. K., Boruah, A. N., Biswas, S. K., Chakraborty, M., & Purkayastha, B. (2022). Dengue Fever Prediction using Machine Learning Analytics. In 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing. <https://doi.org/10.1109/COM-IT-CON54601.2022.9850923>
- Salim, N. A. M., Wah, Y. B., Reeves, C., Smith, M., Yaacob, W. F. W., Mudin, R. N., Dapari, R., Sapri, N. N. F. F., & Haque, U. (2021). Prediction of dengue outbreak in Selangor Malaysia using machine learning techniques. *Scientific Reports*, 11(1), 1–9. <https://doi.org/10.1038/s41598-020-79193-2>
- Sarma, D., Hossain, S., Mitra, T., Bhuiya, M. A. M., Saha, I., & Chakma, R. (2020). Dengue Prediction using Machine Learning Algorithms. In IEEE Region 10 Humanitarian Technology Conference, R10-HTC. <https://doi.org/10.1109/R10-HTC49770.2020.9357035>
- Siddiq, A., Shukla, N., & Pradhan, B. (2021). Predicting Dengue Fever Transmission Using Machine Learning Methods. In 2021 IEEE International Conference on Industrial Engineering and Engineering Management, IEEM 2021, 2007, (21–26). <https://doi.org/10.1109/IEEM50564.2021.9672977>
- Singh, K. D. (2021). Particle Swarm Optimization assisted Support Vector Machine based Diagnostic System for Dengue prediction at the early stage. In Proceedings - 2021 3rd International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2021, (844–848). <https://doi.org/10.1109/ICAC3N53548.2021.9725670>
- Souza, C., Maia, P., Stolerman, L. M., Rolla, V., & Velho, L. (2022). Predicting dengue outbreaks in Brazil with manifold learning on climate data (Formula presented). *Expert Systems with Applications*, 192(November 2020), 116324. <https://doi.org/10.1016/j.eswa.2021.116324>
- Srivastava, S., Soman, S., Rai, A., & Cheema, A. S. (2020). An online learning approach for dengue fever classification. In Proceedings - IEEE Symposium on Computer-Based Medical Systems, (pp. 163–168). <https://doi.org/10.1109/CBMS49503.2020.00038>
- Sukama, Y., Hertono, G. F., Handari, B. D., & Aldila, D. (2020). Comparing activation functions in predicting dengue hemorrhagic fever cases in DKI Jakarta using u. In AIP Conference Proceedings, 2296. <https://doi.org/10.1063/5.0030456>
- Tanawi, I. N., Vito, V., Sarwinda, D., Tasman, H., & Hertono, G. F. (2021). Support Vector Regression for Predicting the Number of Dengue Incidents in DKI Jakarta. *Procedia Computer Science*, 179(2020), 747–753. <https://doi.org/10.1016/j.procs.2021.01.063>

- Xu, J., Xu, K., Li, Z., Meng, F., Tu, T., Xu, L., & Liu, Q. (2020). Forecast of dengue cases in 20 chinese cities based on the deep learning method. *International Journal of Environmental Research and Public Health*, 17(2). <https://doi.org/10.3390/ijerph17020453>
- Yavari Nejad, F., & Varathan, K. D. (2021). Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction. *BMC Medical Informatics and Decision Making*, 21(1), 1–12. <https://doi.org/10.1186/s12911-021-01493-y>