



Universidad
Señor de Sipán

**FACULTAD DE INGENIERIA, ARQUITECTURA Y
URBANISMO**

ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS

TESIS

**Análisis comparativo de algoritmos de aprendizaje
automático para clasificar Dengue Virus**

**PARA OPTAR EL TÍTULO PROFESIONAL DE INGENIERA
DE SISTEMAS**

Autora

Bach. Exebio Chepe Yolanda Victoria de Fatima

ORCID <https://orcid.org/0000-0002-2076-2120>

Asesor

Mg. Bravo Ruiz Jaime Arturo

ORCID <https://orcid.org/0000-0003-1929-3969>

Línea de Investigación

**Ciencias de la información como herramientas multidisciplinares
y estratégicas en el contexto industrial y de organizaciones**

Sublínea de Investigación

**Informática y transformación digital en el contexto industrial y
organizacional**

Pimentel – Perú

2024

**ANALISIS COMPARATIVO DE ALGORITMOS DE APRENDIZAJE AUTOMATICO
PARA CLASIFICAR DENGUE VIRUS**

Aprobación del jurado

MG. BRAVO RUIZ JAIME ARTURO

Presidente del Jurado de Tesis

MG. PINEDO SIFUENTES GIANMARCO JAVIER

Secretario del Jurado de Tesis

MG. BANCES SAAVEDRA DAVID ENRIQUE

Vocal del Jurado de Tesis



Universidad
Señor de Sipán

ANEXO 01: DECLARACIÓN JURADA DE ORIGINALIDAD

Quien(es) suscribe(n) la DECLARACIÓN JURADA, soy(somos) **Yolanda Victoria de Fatima Exebio Chepe** del Programa de Estudios de **Ingeniería de Sistemas** de la Universidad Señor de Sipán S.A.C, declaro bajo juramento que soy autor del trabajo titulado:

ANÁLISIS COMPARATIVO DE ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PARA CLASIFICAR DENGUE VIRUS

El texto de mi trabajo de investigación responde y respeta lo indicado en el Código de Ética del Comité Institucional de Ética en Investigación de la Universidad Señor de Sipán, conforme a los principios y lineamientos detallados en dicho documento, en relación con las citas y referencias bibliográficas, respetando el derecho de propiedad intelectual, por lo cual informo que la investigación cumple con ser inédito, original y autentico.

En virtud de lo antes mencionado, firman:

Exebio Chepe Yolanda Victoria de Fatima	DNI: 71430121	
---	---------------	--

Pimentel, 19 de Diciembre de 2023

Dedicatoria

A mis padres, Alfredo y Fresia cuyo amor y sacrificio han sido la luz que me ha guiado en este largo viaje académico. Su constante apoyo, comprensión y paciencia han sido mi mayor fortaleza. A mis hermanas, sobrina y amigos, quienes han estado a mi lado, celebrando mis éxitos y brindándome consuelo en los momentos difíciles. A cada persona que ha cruzado mi camino y ha dejado una huella en mi corazón, gracias. Este logro es el resultado del amor, la dedicación y el esfuerzo de todos ustedes. Con profunda gratitud, dedico este trabajo a cada uno de ustedes.

La autora

Agradecimientos

A la Universidad Señor de Sipán – Chiclayo, por su inestimable apoyo y por brindarme las herramientas necesarias para alcanzar mis metas académicas.

Al Mg. Jaime Bravo Ruiz, por su apoyo y asesoramiento constante en la investigación desarrollada.

A los servidores de Hospital Almanzor Aguinaga Asenjo de la Red EsSalud, por brindarme su apoyo, para el desarrollo del presente trabajo.

La autora

Índice de Contenidos

DEDICATORIA.....	IV
AGRADECIMIENTOS	V
INDICE DE TABLAS.....	VII
INDICE DE FIGURAS	VII
RESUMEN.....	IX
ABSTRACT	X
I. INTRODUCCIÓN	1
II. MATERIALES Y MÉTODO	4
2.1. MATERIALES	4
2.2. MÉTODO.....	5
III. RESULTADOS Y DISCUSIÓN	18
3.1. RESULTADOS.....	18
3.2. DISCUSIÓN.....	21
IV. CONCLUSIONES Y RECOMENDACIONES	22
4.1. CONCLUSIONES	22
4.2. RECOMENDACIONES.....	23
REFERENCIAS	24
ANEXOS:	29

Índice de Tablas

TABLA I: LISTA DE MATERIALES	4
TABLA II: ESTRUCTURA DEL DATASET	7
TABLA III: ALGORITMOS CON RENDIMIENTOS DE REVISIÓN SISTEMÁTICA.....	8
TABLA IV: ALGORITMOS SELECCIONADOS.....	9
TABLA V: MEDIDA DE CONSUMO DE CPU, MEMORIA RAM Y TIEMPO DE RESPUESTA UTILIZADOS POR LOS TRES ALGORITMOS DE APRENDIZAJE AUTOMATICO.....	18
TABLA VI: RENDIMIENTO DE LOS TRES ALGORITMOS DE APRENDIZAJE AUTOMÁTICO	19

Índice de Figuras

FIG. 1. DIAGRAMA DEL MÉTODO UTILIZADO. ELABORACIÓN PROPIA	5
FIG. 2.CARGA DE LA DATASET.....	11
FIG. 3.TRANSFORMACIÓN DE VARIABLES	11
FIG. 4.DIVISIÓN DE DATA	12
FIG. 5.CREACIÓN DEL MODELO SVM LINEAL.....	12
FIG. 6.EVALUACIÓN DE METRICAS	13
FIG. 7.ARQUITECTURA DE RANDOM FOREST.	13
FIG. 8.GENERADOR DEL MODELO RF.....	14
FIG. 9.EVALUACIÓN DE METRICAS	14
FIG. 10.GENERACIÓN EL ARBOL.....	15
<i>FIG. 11. RANDOM FOREST GENERADO EN LA INVESTIGACIÓN.</i>	<i>15</i>
FIG. 12.REDES NEURONALES ARTIFICIALES	16
FIG. 13.ARQUITECTURA DE UNA RED NEURONAL ARTIFICIAL	16
FIG. 14.CONSTRUCCIÓN DEL MODELO	17
FIG. 15.EVALUACIÓN DE METRICAS	18

FIG. 16.MEDIDAS DE CONSUMO DE CPU, MEMORIA RAM Y TIEMPO DE RESPUESTA UTILIZADA POR LOS TRES ALGORITMOS DE APRENDIZAJE AUTOMATICO..... 19

FIG. 17.EVALUACIÓN DE RENDIMIENTO DE LOS ALGORITMOS DE APRENDIZAJE AUTOMÁTICO PLANTEADOS EN NUESTRA INVESTIGACIÓN 20

Resumen

La enfermedad del dengue, propagada por el mosquito *Aedes aegypti*, requiere una clasificación precisa de casos para gestionarla de manera efectiva. En este ámbito, la aplicación de algoritmos de aprendizaje automático surge como una herramienta que puede mejorar el diagnóstico y contribuir a estrategias de control más eficientes. Por esta razón, la investigación se enfoca en realizar un análisis comparativo de algoritmos de aprendizaje automático para clasificar el virus del dengue. Se empleó un conjunto de datos de pacientes de un hospital público, recopilando 21,157 datos bajo los criterios de periodo, resultados, sexo, edades, síntomas, autóctono e importado. A través de una revisión de la literatura, se seleccionaron los algoritmos, entre ellos Support Vector Machine, Random Forest y Artificial Neural Network, que demostraron los mejores resultados en precisión. Posteriormente, se fraccionó el conjunto de datos en 14,809 datos (70%) para el entrenamiento y 6,348 datos (30%) para las pruebas. La evaluación del desempeño se llevó a cabo utilizando métricas como precisión, recall y especificidad. Los resultados indicaron que la Red Neuronal Artificial lidera con una precisión del 86.47% y un recall del 92.91% en la clasificación de casos relacionados con el dengue. Al evaluar tres algoritmos de clasificación, la Support Vector Machine destaca por su alto recall, Random Forest presenta un equilibrio sólido, y la Artificial Neural Network exhibe la mayor precisión.

Palabras Clave: Aprendizaje automatico, salud, algoritmos, metricas, dengue

Abstract

Dengue disease, spread by the *Aedes aegypti* mosquito, requires accurate classification of cases to manage it effectively. In this area, the application of machine learning algorithms emerges as a tool that can improve diagnosis and contribute to more efficient control strategies. For this reason, the research focuses on performing a comparative analysis of machine learning algorithms to classify dengue virus. A dataset of patients from a public hospital was used, collecting 21,157 data under the criteria of period, outcome, sex, age, symptoms, autochthonous and imported. Through a review of the literature, algorithms were selected, including Support Vector Machine, Random Forest and Artificial Neural Network, which showed the best results in accuracy. Subsequently, the data set was split into 14,809 data (70%) for training and 6,348 data (30%) for testing. Performance evaluation was carried out using metrics such as accuracy, recall and specificity. The results indicated that the Artificial Neural Network leads with an accuracy of 86.47% and recall of 92.91% in the classification of dengue-related cases. When evaluating three classification algorithms, the Support Vector Machine stands out for its high recall, Random Forest presents a solid balance, and the Artificial Neural Network exhibits the highest accuracy.

Keywords: Machine learning, health, algorithms, metrics, dengue fever

I. INTRODUCCIÓN

El dengue es una enfermedad ocasionada por flavivirus que es transmitida a través de los mosquitos *Aedes aegypti*. Es muy común en más de 125 países en área tropicales y subtropicales, al año causa alrededor de trescientos noventa millones de contagios a nivel mundial, de las cuales 96 millones han presentado síntomas clínicos. El riesgo que contraer la enfermedad está presente para cualquier persona que viva o viaje a una región donde el dengue es endémico, como lo son: Asia tropical, América Central y del Sur y El Caribe. Por lo general la transmisión en estas regiones ocurren todo el año.[1]

A pesar que el dengue es conocida como una enfermedad tropical, se notificaron casos originalmente en el sur de los EE.UU. y algunos países europeos como Italia, Francia, Grecia y España.[2]

El diagnóstico del dengue se lleva a cabo a través de la unión de técnicas de laboratorio, como la Polymerase Chain Reaction with Reverse Transcription (PCR-RT) específica para el dengue, la evaluación de la presencia de la proteína "Non-Structural" (proteína NS1), la detección de Inmunoglobulina tipo M o el análisis de incremento en las diluciones de la Inmunoglobulina tipo G, así como el aislamiento del virus. Puede considerarse como un enfoque confirmatorio cualquier combinación de estas pruebas. Además, los síntomas y signos clínicos desempeñan un papel importante en la diferenciación diagnóstica de la infección viral del dengue. [3]

El propósito de esta investigación es realizar un análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus. Esto puede ser crucial en la gestión de la salud pública, pues el Dengue es una enfermedad que impacta a millones de personas a nivel global. En el campo de la salud el uso de algoritmos de aprendizaje automático es utilizado dado que cuentan con la capacidad de analizar datos complejos sin restricciones de métodos de modelado tradicionales. En el caso del dengue, los algoritmos se utilizan en estudios de diagnósticos intervención o pronóstico de epidemias. [4]

A partir de ello surgió la siguiente pregunta de investigación ¿Cuál es el algoritmo de aprendizaje automático que presenta mejor desempeño en la clasificación de Dengue Virus?,

esto conlleva a plantear la siguiente hipótesis, que el algoritmo de aprendizaje automático que presenta mejor desempeño en la clasificación de Dengue Virus es Support Vector Machine.

Realizar la investigación, es de gran alcance debido a que incrementa la exactitud del diagnóstico y el tratamiento oportuno; lo que permite optimizar el uso de los recursos asignados a salud; identificando los algoritmos más eficientes, ayuda a identificar características y factores de riesgo asociados al dengue, facilitando el dictamen clínico y contribuye al avance científico en el ámbito antes mencionado.

Shaikh et al. [5] indican que el problema que abarca es sobre la predicción temprana de la etapa de escape plasmático en pacientes con dengue, se utilizan algoritmos como Random Forest y Light Gradient Boosting Machine (LightGBM), como resultados se obtuvieron 76,9% de precisión y un 87,9% de especificidad en el conjunto de prueba, concluyen que beneficiaría en la toma de decisiones para la admisión hospitalaria de casos de sospecha de dengue. También, Yang et al.[6] mencionan que identifican las principales causas que soportan la aparición de esta condición a través de algoritmos de machine learning como el árbol de clasificación y regresión (CART) y Random Forest (RF), estos resultados revelan que el modelo RF identificó patrones altamente complejos al examinar cómo las características demográficas, los síntomas clínicos y los marcadores bioquímicos interactúan entre sí, permitiendo prever el 43,1% de los casos de dengue grave, concluyen que esto es beneficioso para identificar temprano casos graves de dengue. Además, Handari et al. [7] comentan que buscan saber qué modelo y distribución de datos producen la mejor predicción de casos de Dengue Hemorrágico, para ello utilizaron tres modelos de machine learning y diferentes divisiones de datos. El modelo Long Short-Term Memory (LSTM) sobresalió con un error cuadrático medio de 0,0442, ofreciendo las mejores predicciones de casos en DKI Jakarta. Estos resultados son valiosos para las autoridades sanitarias en la prevención y control de futuros brotes. Asimismo, Nejad et al. [8] manifiestan que quieren predecir brotes de dengue y mejorar la exactitud con aprendizaje automático. Evaluaron cinco modelos, destacando Bayes network con nuevo factor de riesgo, alcanzando 92.35% de

precisión en la predicción. Concluyen que identificando factores mejoran la predicción de brotes de dengue, de igual manera, Lim et al. [9] explican que desean predecir epidemias de dengue mediante modelos bayesianos de cambio de régimen, empleando algoritmos como BAR y BRS. El modelo BRS demostró una mayor capacidad predictiva que los modelos de series temporales, identificando la influencia de factores climáticos en la transmisión del dengue, con un MAPE del 5,43%. Concluyen que los modelos de cambio de régimen fueron más adecuados para comprender las dinámicas de la transmisión del dengue, sugiriendo su valor en epidemiología. Asimismo, Li et al. [10] dicen que buscan mejorar la precisión de las predicciones de casos de dengue con datos de series de tiempo limitados. Utilizaron la plataforma Earth Engine (GEE) y el modelo LSTM para lograrlo, obteniendo un bajo error cuadrático medio (0.33%) y error porcentual absoluto medio (0.28%) en la semana 12. Este enfoque tiene un gran potencial para pronosticar el riesgo de dengue de manera efectiva y práctica, y se sugiere su aplicación en otras regiones o a nivel global para fortalecer las estrategias de prevenir y controlar la enfermedad, también, Tanawi et al. [11] señalan que se enfocan en predecir casos de dengue en DKI Jakarta usando datos climáticos y previos. Se compararon modelos de regresión de vectores de soporte con kernel lineal y radial, siendo el lineal más preciso con un error cuadrático medio del 3.6678%. Variables como temperatura y lluvia acumulada correlacionaron positivamente, mientras que la temperatura media lo hizo negativamente, destacando la influencia del clima en los casos de dengue, además, Kaur et al. [12] sostienen que desean predecir y diagnosticar el dengue utilizando diversos algoritmos de aprendizaje automático en Python. Donde DT, CART y NB mostraron mayor precisión, sensibilidad, especificidad y recuperación. Desarrollaron un modelo que predice la infección por dengue basado en signos vitales y síntomas, con una precisión superior al 90% y una tasa de clasificación precisa del 90%. Los algoritmos de aprendizaje automático extraen patrones ocultos en los datos del dengue para una detección precisa y mejorar la gestión de la enfermedad, finalmente, Hamdanie et al. [13] expresan que buscan clasificar pacientes con dengue utilizando aprendizaje automático y síntomas como datos, compararon algoritmos como KNN, RF, DT y SVM, encontrando que SVM fue el más efectivo con una precisión,

recuperación y exactitud del 99.1%. Concluyeron que el aprendizaje automático puede ser altamente efectivo para diagnosticar y predecir el tipo de dengue basado en síntomas, lo que es crucial para un tratamiento adecuado y una detección temprana.

II. Materiales y Método

2.1. Materiales

TABLA I:
LISTA DE MATERIALES

Ítem	Nombre	Descripción
1	Data set de personas que han presentado síntomas de dengue	Consta de 21157 datos clínicos Datos reales
2	Software	Google Colab
3	Laptop	Procesador: 11th Gen Intel(R) Core(TM) i5-1135G7 2.40GHz RAM: 8.00 GB Gráficos Intel® Iris® Xe e Iris® Sistema operativo Windows 10
4	Lenguaje Python	Cuenta con librerías especializadas en aprendizaje automático Se destaca por su sintaxis clara y legible

Nota: Listado de materiales para la investigación. Elaboración propia

2.2. Método



Fig. 1. Diagrama del método utilizado. Elaboración Propia

Presentan el método utilizado en la investigación en la figura 1, en la primera etapa, se construye la dataset sobre casos de personas han presentado síntomas del dengue, la segunda etapa, se seleccionaron los algoritmos que han presentado mejores resultados en cuanto a los criterios establecidos, posteriormente se entrenaron los algoritmos utilizando el entorno Google Colab utilizando Python como lenguaje de programación, seguido de ello se evaluaron los resultados de las diferentes métricas.

En una primera instancia fue necesario construir una dataset con la cual se deben implementar los algoritmos de aprendizaje automático. Se consideró que era esencial recopilar información de fuentes confiables en el departamento de Lambayeque, específicamente en el campo de la salud pública. Se enfocó en recolectar datos reales y relevantes sobre casos de dengue que se han reportado en los primeros meses del año. Se siguió ese camino debido a que, si bien existen dataset relacionados con el dengue, disponibles en plataformas como Kaggle y otros, muchas de ellas están vinculadas a datos

que no son relevantes para esta investigación. Sin embargo, el interés radica en trabajar con datos clínicos específicos dentro de la región Lambayeque.

Se obtuvo los datos de un hospital público siguiendo los protocolos de confiabilidad y seguridad de la información personal de los individuos involucrados. Esto permitió recopilar una base de datos confiable y valiosa que sirvió para la investigación.

Para la creación de la dataset, se siguió una serie de criterios rigurosos para asegurar que la información recopilada sea precisa y relevante. A continuación, se describirá estos criterios detalladamente:

Período: Se refiere a datos recopilados durante los meses de enero a julio de 2023. Esta elección se basa en la observación donde se evidenció que durante este período se registró un notable aumento en la cifra de casos de dengue.

Resultado: Los datos utilizados en este campo se dividen en dos categorías principales: "confirmado" y "descartado". Estos resultados se obtienen a partir de las pruebas realizadas a los pacientes, lo que permitió clasificar con precisión los casos de dengue.

Edad: Se clasificó a las personas en grupos etarios siguiendo la normativa técnica establecida por el Ministerio de Salud en el Perú. Estos grupos son:

Niño: 0 a 11 años 11 meses y 29 días

Adolescente: 12 a 17 años 11 meses y 29 días

Joven: 18 a 29 años 11 meses y 29 días

Adulto: 30 a 59 años 11 meses y 29 días

Adulto mayor: 60 años o más

Sexo: Se clasificó los casos de dengue en dos categorías: Femenino y Masculino, con el fin de analizar las diferencias en la incidencia del dengue en función del género.

Síntomas: Se ha tenido en cuenta síntomas relacionados con el dengue. Estos síntomas pueden variar en gravedad, desde leves hasta causar daños significativos o incluso amenazar la vida en algunos casos. Los síntomas típicos del dengue suelen manifestarse entre 4 y 10 días después de ser picado por un mosquito portador de infecciones, teniendo en cuenta la presencia de fiebre, dolor de cabeza, dolor de ojos, dolor de huesos, náuseas y

sarpullido.

Importados: Se ha considerado casos de dengue que se contrajeron fuera de la región de Lambayeque, lo que nos permite analizar la propagación del virus desde otras áreas.

Autóctonos: Se ha considerado casos de dengue que se originaron dentro de la región de Lambayeque. Esto es esencial para comprender la dinámica local de la enfermedad y su impacto en la comunidad.

Después de definir con claridad los criterios para la creación de nuestra dataset, se llevó a cabo la selección y recopilación de información relevante. El resultado de este proceso fue un conjunto de datos sólido y representativo que consta de un total de 21,157 casos, incluyendo tanto aquellos confirmados como descartados.

Posteriormente, se convirtió en un formato .csv que es ampliamente utilizado en análisis de datos y aprendizaje automático. Esto permitirá entrenar nuestros modelos de manera adecuada y realizar análisis más avanzados.

**TABLA II:
ESTRUCTURA DEL DATASET**

Peri odo	Resulta do	Edad	Sexo	Síntomas						Importa do	Autócto no
				Fiebre	Dolor Cabeza	Dolor Ojos	Dolor Huesos	Nause as	Sarpu llido		
Ene ro	confirma do	18_29	Feme nino	Si	Si	Si	Si	No	Si	Si	No
Ene ro	descarta do	0_11	Feme nino	Si	Si	Si	No	Si	No	No	Si
Ene ro	descarta do	mayor 60	Feme nino	No	No	Si	Si	No	Si	No	Si

Nota: Estructura del dataset elaborado para la investigación

Un paso importante de investigación, es la realización de una revisión sistemática de la literatura en el campo de los algoritmos de aprendizaje automático. Este exhaustivo análisis tiene como propósito principal brindarnos una visión actualizada del estado de la investigación en el apasionante dominio del aprendizaje automático.

Esta revisión se transforma en una herramienta esencial para identificar los algoritmos que poseen una relevancia significativa y que se ajustan de manera adecuada a la resolución

de problemas específicos en diversos ámbitos.

Un componente clave de la revisión implica la búsqueda y análisis de estudios publicados en reconocidas bases de datos científicas, tales como Scopus, PubMed y Science Direct.

Este enfoque le proporciona información altamente valiosa sobre cuáles algoritmos destacan como los más efectivos para abordar tareas específicas, permitiéndole tomar decisiones más informadas en proyectos futuros de aprendizaje automático.

Se ha definido con precisión una serie de criterios fundamentales que guiarán nuestra exploración y evaluación en el campo del aprendizaje automático aplicado a la clasificación del dengue. Estos criterios son esenciales para enfocarse en aspectos clave de interés y obtener resultados significativos. Los criterios establecidos son los siguientes:

Algoritmos utilizados en la clasificación de dengue

Evaluar el rendimiento de las métricas que utilizan en el desempeño de los algoritmos de aprendizaje para la clasificación de dengue

Identificar el algoritmo que es más eficiente en términos de tiempo de entrenamiento.

**TABLA III:
ALGORITMOS CON RENDIMIENTOS DE REVISIÓN SISTEMÁTICA**

Artículo	Algoritmo	Precisión	Exactitud	Recall	Sensibilidad	Especificidad
[10], [14]– [17]	Decision Tree (DT)	79%	80%	81%		
[12], [18], [19]	K-Nearest-Neighbor (KNN)	80%	85%			
[6], [9], [16], [17], [20]– [25]	Random Forest (RF)	86.18%	99,1%	99,1%	90%	90%
[5], [8], [13], [14], [16], [17], [19],	Support Vector Machine (SVM)	90.90%	98%	98%	90%	90%

[23], [24], [26]–[31]					
[16], [20], [23], [31]	Logstic Regression (LR)	82%	88%	88%	
[23], [31], [32]	Naive Bayes (NB)	83%	95%	95%	
[28], [31], [33], [34]	Long Short- Term Memory (LSTM)	83%	96%		
[28], [30], [35]–[37]	Redes neuronales convolucionales (CNN)	84,5%			
[5], [23]–[26], [28], [29], [32]	Artificial Neural Networks (ANN)	94.38%	97.8%	96%	97%
[7], [18], [35], [38]	Recurrent Neural Networks (RNN)	76%			
[12], [28], [39]	Linear Discriminant Analysis (LDA)	82%			

Nota: Rendimiento de los algoritmos encontrados en la revisión sistemática

En base en los criterios previamente mencionado, se elaboró una lista de algoritmos seleccionados, con el propósito de entrenarlos y, de este modo, comparar los resultados obtenidos.

TABLA IV:
ALGORITMOS SELECCIONADOS

N°	Algoritmo	Precisión
1	Random Forest (RF)	86.18%
2	Support Vector Machine (SVM)	90.90%

Nota: Algoritmos que han presentado mejores resultados

Las métricas seleccionadas para esta evaluación de algoritmos son las siguientes:

Precisión: Esta métrica desempeña un papel crucial al permitir conocer el porcentaje de valores que han sido clasificados como positivos y que realmente son positivos. En un contexto médico, la precisión proporciona una visión clara de que tan acertadamente se detectan los casos reales de la enfermedad entre todos los casos catalogados como positivos.

Recall: Permite saber cuántos valores positivos son clasificados correctamente. En el ámbito de la salud, determina la capacidad del algoritmo para detectar la enfermedad entre los individuos enfermos.

Especificidad: Permite saber cuántos valores negativos son clasificados correctamente. En un contexto médico, se refiere a la capacidad de identificar correctamente los casos de pacientes sanos entre el grupo de individuos que son efectivamente sanos. Minimiza los falsos positivos.

Para entrenar los algoritmos, se empleó el entorno Google Colab, el cual no requiere la instalación ni actualización del hardware para cumplir con los requisitos intensivos de CPU/GPU de Python. Proporciona acceso gratuito a diversos recursos informáticos, como almacenamiento, memoria, capacidad de procesamiento, unidades de procesamiento gráfico (GPU) y unidades de procesamiento tensorial (TPU). Además, utiliza el lenguaje de programación Python.

Como se mencionó previamente, se seleccionaron los algoritmos SVM, RF y ANN.

a. Support Vector Machine

Se inicia con la carga de la data que contiene información sobre distintos casos de dengue, incluyendo el periodo, resultado, edad, sexo, síntomas, importado y autóctono.

```

# Cargo los datos del csv
#ruta = "dataset.csv"
ruta = "dataset.csv"
data = pd.read_csv(ruta, sep=',')
data.head(5)

```

	periodo	resultado	edad	sexo	fiebre	dolorCabeza	dolorOjos	dolorHuesos	nauseas	sarpullido	importado	autoctono
0	1	1	3	1	1	1	1	1	0	1	1	0
1	1	0	1	1	1	1	1	0	1	0	0	1
2	1	0	5	1	0	0	1	1	0	1	0	1
3	1	0	5	0	1	0	0	0	0	0	1	0
4	1	0	4	0	1	1	0	1	1	1	1	0

Fig. 2.Carga de la dataset

Luego se realiza la transformación de variables categóricas y binarias a variables numéricas utilizando el método fit_transform de LabelEncoder (le). Esta etapa es esencial para permitir que la Máquina de Soporte Vectorial pueda operar con dichas variables.

```

### Recategorizo variables
# Periodo puede que la deje afuera, no aporta mucho
data['periodo'] = le.fit_transform(data['periodo'])
data['edad'] = le.fit_transform(data['edad'])
data['sexo'] = le.fit_transform(data['sexo'])
data['fiebre'] = le.fit_transform(data['fiebre']) # numerica
data['dolorCabeza'] = le.fit_transform(data['dolorCabeza'])
data['dolorOjos'] = le.fit_transform(data['dolorOjos'])
data['dolorHuesos'] = le.fit_transform(data['dolorHuesos'])
data['nauseas'] = le.fit_transform(data['nauseas'])
data['sarpullido'] = le.fit_transform(data['sarpullido'])
data['importado'] = le.fit_transform(data['importado'])
data['autoctono'] = le.fit_transform(data['autoctono'])

```

Fig. 3.Transformación de variables

Se definen las variables predictoras y la variable objetivo para el modelo. El conjunto de variables predictoras, denominado como X, está compuesto por diversas características, tales como 'periodo', 'edad', 'sexo', 'fiebre', 'dolorCabeza', 'dolorOjos', 'dolorHuesos', 'nauseas', 'sarpullido', 'importado' y 'autoctono'. Estas características representan los atributos que se utilizarán para realizar predicciones sobre el resultado del modelo. Por otro lado, la variable objetivo, representada por y, está asociada a la columna 'resultado'. Esta variable contiene las etiquetas de clase que el modelo SVM buscará predecir durante el proceso de entrenamiento y evaluación.

Luego, se lleva a cabo la partición del conjunto de datos en grupos de entrenamiento y prueba mediante el uso de la función `train_test_split`. Es importante destacar que se asigna un 30% del conjunto de datos para el conjunto de prueba, mientras que el 70% se destina al conjunto de entrenamiento. Esta proporción es determinada por el parámetro `test_size=0.3`. Además, se utiliza el parámetro `random_state=2023` con el propósito de establecer una semilla aleatoria y así garantizar la reproducibilidad del experimento en futuras ejecuciones..

```
# Generando la muestra de prueba y la muestra de entrenamiento
# variables predictoras
X = data[['periodo', 'edad', 'sexo', 'fiebre',
         'dolorCabeza', 'dolorOjos', 'dolorHuesos',
         'nauseas', 'sarpullido', 'importado', 'autoctono']]
y = data['resultado'] # variable objetivo
#Separar en conjunto de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=2023)
# imprimimos la longitud de cada conjunto
print("Datos de entrenamiento:", len(X_train))
print("Datos de prueba:", len(X_test))
```

Fig. 4. División de data

A continuación, se crea una SVM con un kernel lineal utilizando la clase `SVC()`. El modelo SVM se entrena utilizando el conjunto de entrenamiento mediante la instrucción `svm.fit(X_train, y_train)`. La evaluación del modelo se lleva a cabo con el conjunto de prueba (`X_test`), donde la SVM clasifica los datos no observados y se cotejan las predicciones con las etiquetas reales (`y_test`).

```
[ ] # ## Modelo SVM lineal
    ### Generando la maquina de soporte vectorial con kernel lineal SVM ###
    svm = SVC()
    svm.fit(X_train, y_train)
```

Fig. 5. Creación del modelo SVM Lineal

Se calculan diversas métricas de evaluación, como precisión, recall y especificidad, proporcionando datos exhaustivos acerca del desempeño del modelo en cuanto a la precisión de la clasificación.

```
[ ] ### Evaluando el modelo ###
y_pred = svm.predict(X_test)
start = time.time()
### Precisión del modelo ###
accuracy = accuracy_score(y_test, y_pred)
print("Precisión:", round(accuracy*100 ,2),"%")
### Recall del modelo ###
recall = recall_score(y_test, y_pred)
print("Recall:", round(recall*100 ,2),"%") #
### Especificidad
# Obtener la matriz de confusión
confusion = confusion_matrix(y_test, y_pred)
# Extraer los valores de la matriz de confusión
tn, fp, fn, tp = confusion.ravel()
```

Fig. 6.Evaluación de metricas

b. Random Forest

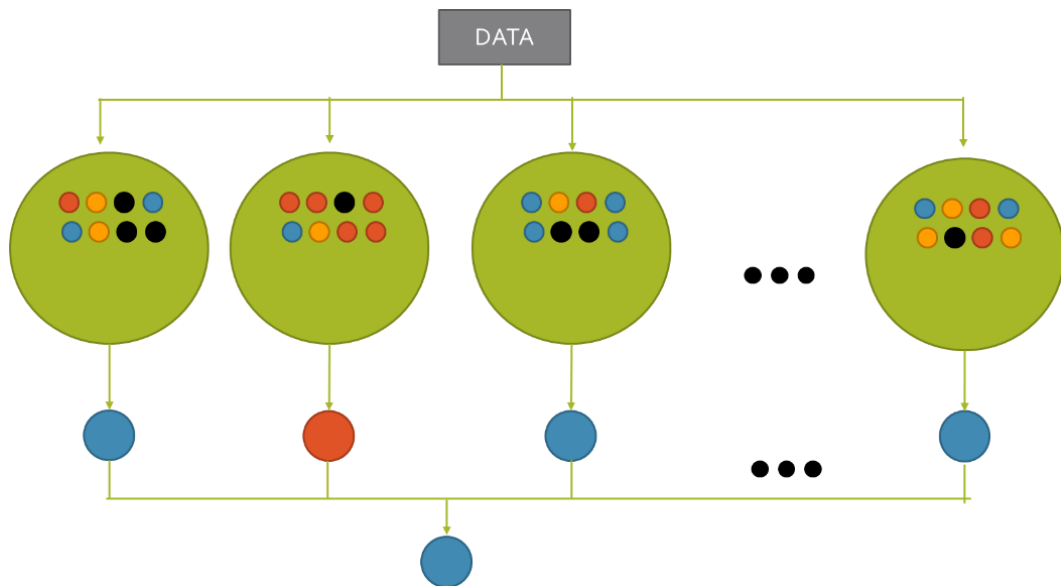


Fig. 7.Arquitectura de Random Forest.

Se inicia con la carga de la data antes mencionado y de igual forma se realiza la transformación de variables categóricas y binarias a variables numéricas. Este procedimiento es esencial para facilitar la manipulación de las variables por parte del modelo Random Forest. Posteriormente, se lleva a cabo la división del conjunto de datos en conjuntos de entrenamiento y prueba mediante la función `train_test_split`.

Seguidamente, se genera un modelo de Random Forest empleando la clase `RandomForestClassifier()`. El modelo se entrena utilizando el conjunto de entrenamiento mediante la instrucción `rf.fit(X_train, y_train)`.

```

# Separar en conjunto de entrenamiento y prueba
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3, random_state=2023)

### Generando el random forest ###
rf = RandomForestClassifier()
rf.fit(X_train, y_train)

```

Fig. 8. Generador del modelo RF

La evaluación del modelo se realiza utilizando el conjunto de prueba (X_{test}), donde RF clasifica los datos no observados y se comparan las predicciones con las etiquetas reales (y_{test}). Diversas métricas de evaluación, como precisión, recall y especificidad, son calculadas para proporcionar una comprensión detallada del rendimiento del modelo

```

[ ] ### Evaluando el modelo ###
y_pred = rf.predict(X_test)
start = time.time()
### Precisión del modelo ###
accuracy = accuracy_score(y_test, y_pred)
print("Precisión:", round(accuracy*100 ,2), "%")
### Recall del modelo ###
recall = recall_score(y_test, y_pred)
print("Recall:", round(recall*100 ,2), "%") #
### Especificidad
# Obtener la matriz de confusión
confusion = confusion_matrix(y_test, y_pred)
# Extraer los valores de la matriz de confusión
tn, fp, fn, tp = confusion.ravel()
# Calcular la especificidad
especificidad = tn / (tn + fp)
print("Especificidad:", round(especificidad*100 ,2), "%")

```

Fig. 9. Evaluación de métricas

Adicionalmente, se lleva a cabo la visualización de un árbol individual mediante `tree.plot_tree`, y la representación gráfica se guarda en un archivo de imagen. Esta práctica permite examinar la estructura y las decisiones tomadas por un árbol específico dentro del conjunto del Random Forest.


```
[ ] # Gráficoando el arbol
from sklearn.tree import export_graphviz
import matplotlib.pyplot as plt
from sklearn import tree
###
fig, axes = plt.subplots(nrows = 1,ncols = 1,figsize = (4,4), dpi=800)
tree.plot_tree(rf.estimators_[0],
              filled = True)
fig.savefig('rf_individualtree.png')
```

Fig. 10. Generación el arbol

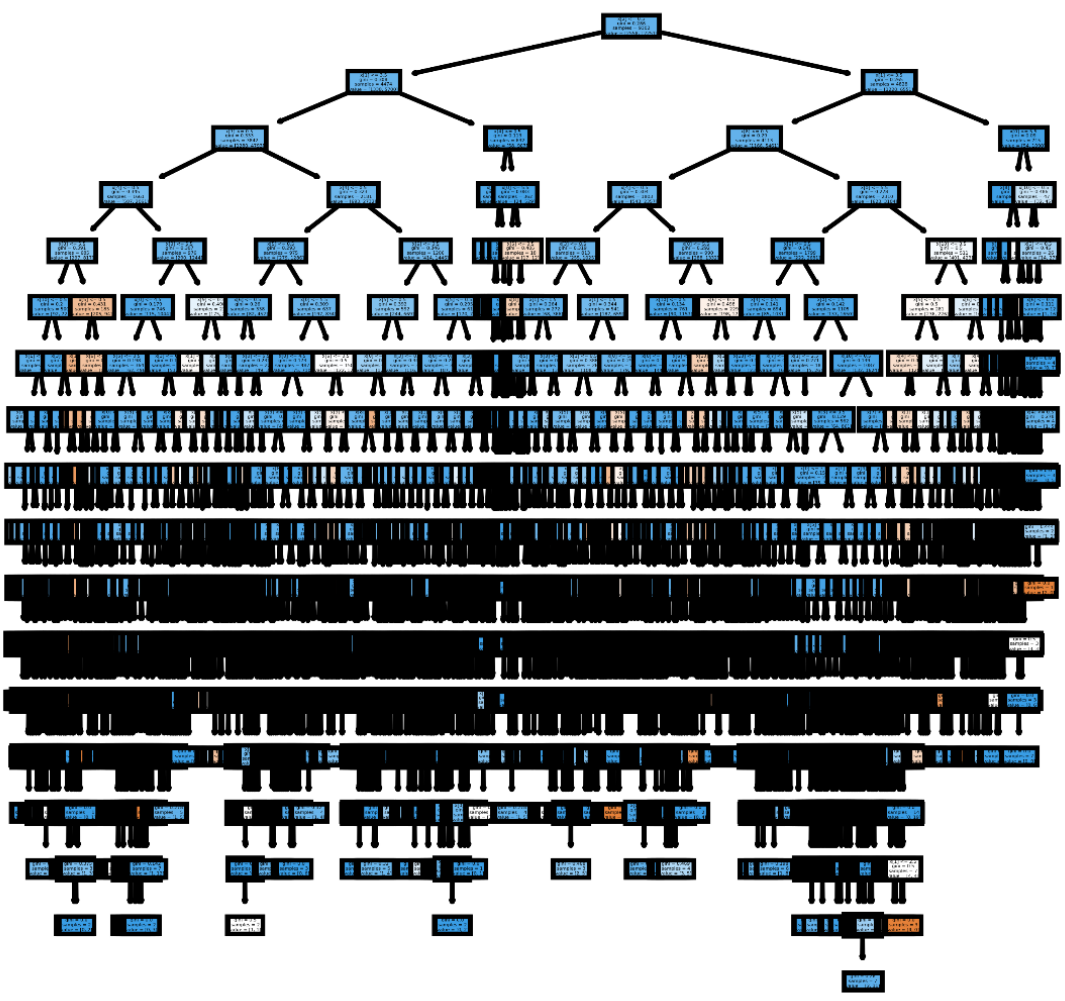


Fig. 11. Random Forest generado en la investigación.

c. Artificial Neural Network

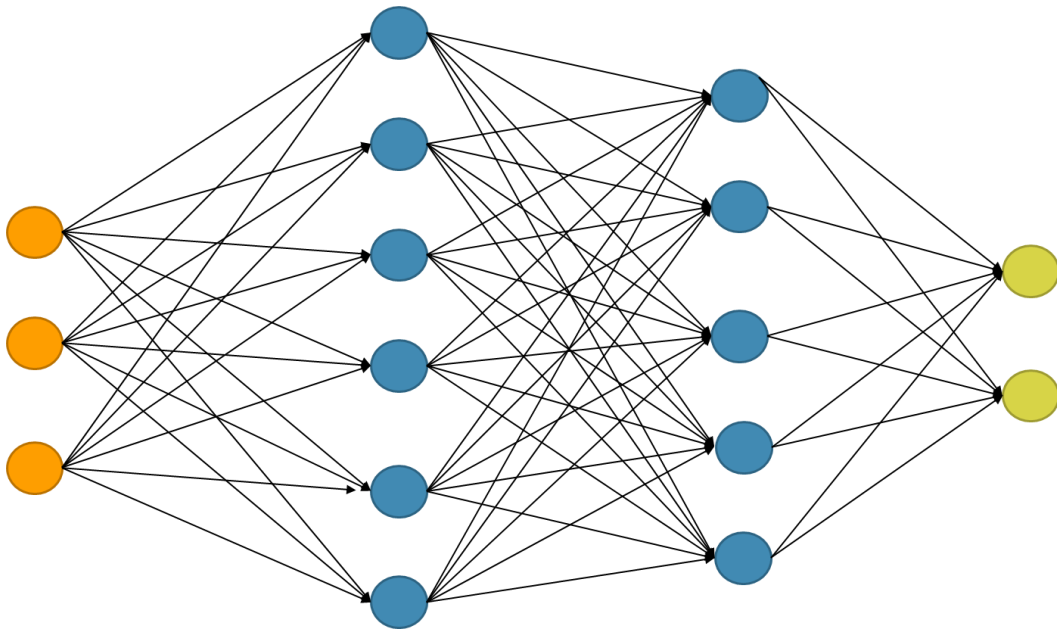


Fig. 12. Redes Neuronales Artificiales

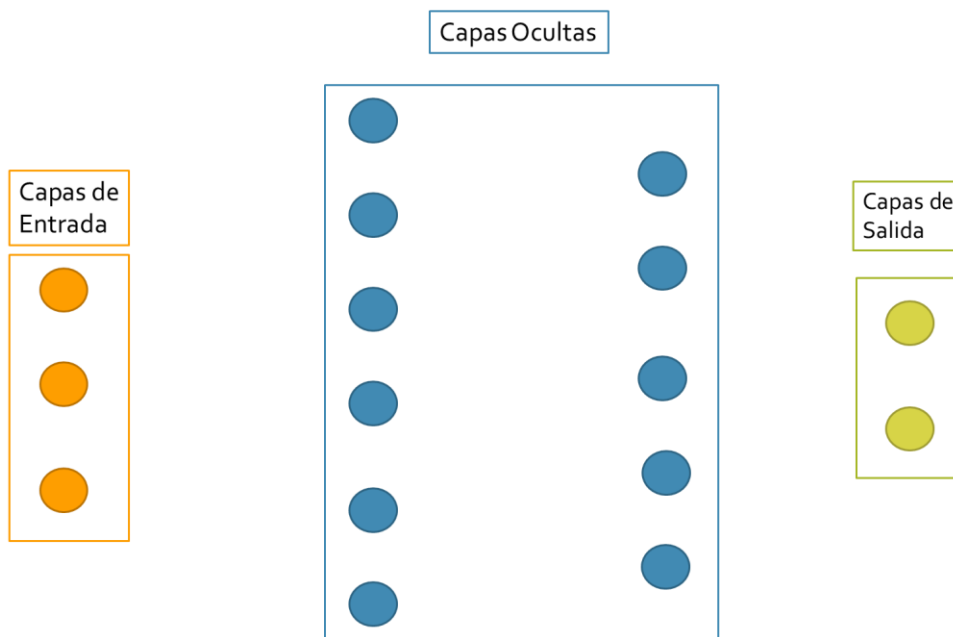


Fig. 13. Arquitectura de una red neuronal artificial

Se inicia con la carga de la data antes mencionado y de igual forma se realiza la transformación de variables categóricas y binarias a variables numéricas. Este procedimiento es fundamental para permitir que una red neuronal artificial procese y aprenda de dichas variables.

A continuación, el conjunto de datos se divide en conjuntos de entrenamiento y prueba a través de la función `train_test_split`, garantizando una evaluación precisa del modelo en datos no vistos durante el proceso de entrenamiento.

Posteriormente, se procede a la construcción y entrenamiento del modelo de redes neuronales Multilayer perceptron mediante la clase `MLPClassifier` de la biblioteca `scikit-learn`. La arquitectura de la red se especifica con dos capas ocultas, una compuesta por 64 neuronas y otra por 32 neuronas, lo que confiere al modelo cierta complejidad para capturar patrones en los datos. El modelo se entrena utilizando el conjunto de entrenamiento mediante el método `fit`.

```
# # Redes neuronales artificiales
import time
import psutil

# Crear el modelo de Red Neuronal
mlp = MLPClassifier(hidden_layer_sizes=(64, 32), max_iter=1000, random_state=42)
# Puedes ajustar el número de capas y neuronas ocultas
# Entrenar el modelo
mlp.fit(X_train, y_train)
```

Fig. 14. Construcción del modelo

En la siguiente etapa, se evalúa la capacidad predictiva del modelo empleando el conjunto de prueba. Se generan predicciones con el método `predict`, y se calculan métricas de evaluación, tales como precisión, recall y especificidad.

Es relevante señalar que la función de activación ReLU se emplea de manera predeterminada en las capas ocultas de la red neuronal. Además, se establece un número máximo de iteraciones (`max_iter`) para el proceso de entrenamiento del modelo

```

# Evaluar el modelo
y_pred = mlp.predict(X_test)
start = time.time()
### Precisión del modelo ###
accuracy = accuracy_score(y_test, y_pred)
print("Precisión:", round(accuracy*100 ,2), "%")
### Recall del modelo ###
recall = recall_score(y_test, y_pred)
print("Recall:", round(recall*100 ,2), "%") #
### Especificidad
# Obtener la matriz de confusión
confusion = confusion_matrix(y_test, y_pred)
# Extraer los valores de la matriz de confusión
tn, fp, fn, tp = confusion.ravel()
# Calcular la especificidad
especificidad = tn / (tn + fp)
print("Especificidad:", round(especificidad*100 ,2), "%")

```

Fig. 15. Evaluación de métricas

III. RESULTADOS Y DISCUSIÓN

3.1. Resultados

TABLA V:

MEDIDA DE CONSUMO DE CPU, MEMORIA RAM Y TIEMPO DE RESPUESTA UTILIZADOS POR LOS TRES ALGORITMOS DE APRENDIZAJE AUTOMÁTICO.

Algoritmo de aprendizaje automático	Consumo de CPU	Consumo de RAM	Tiempo de Respuesta
Support Vector Machine	35.4%	1.06 GB	0.011 seg
Random Forest	21.0%	1 GB	0.024 seg
Artificial Neural Network	62.3%	1.32 GB	0.051 seg

NOTA: Los datos fueron obtenidos teniendo en cuenta las características técnicas del ordenador mencionado anteriormente. Fuente: Elaboración propia

La Tabla 5 indica el consumo de CPU, consumo de memoria RAM y el tiempo de respuesta los cuales fueron necesarios para la ejecución de cada uno de los algoritmos de aprendizaje automático.

El algoritmo SVM utilizó el 35.4% de consumo de CPU, consumió 1.06 GB de RAM, y tuvo un tiempo de respuesta de 0.01 segundos; RF utilizó 21% de consumo de CPU, consumió 1 GB de RAM, y tuvo un tiempo de respuesta de 0.02 segundos; ANN utilizó 62.3%

de consumo de CPU, consumió 1.32 GB de RAM, y tuvo un tiempo de respuesta de 0.05 segundos.

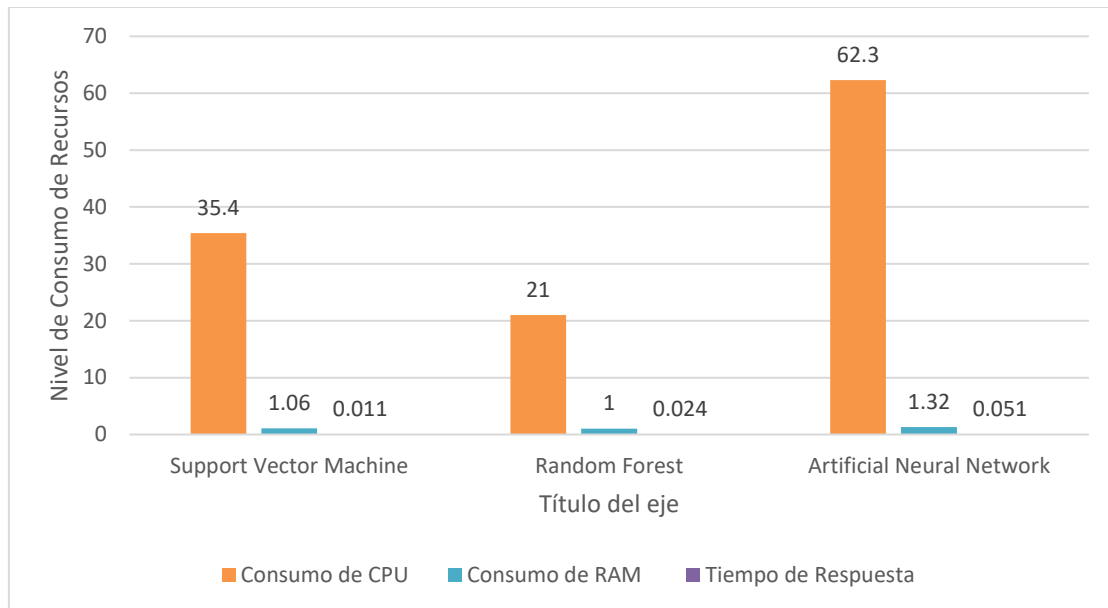


Fig. 16. Medidas de consumo de CPU, memoria RAM y tiempo de respuesta utilizada por los tres algoritmos de aprendizaje automático.

A continuación, se muestran los resultados que se obtuvieron de acuerdo a las métricas de rendimiento de cada algoritmo para clasificar dengue.

TABLA VI:
RENDIMIENTO DE LOS TRES ALGORITMOS DE APRENDIZAJE AUTOMÁTICO

Algoritmo	Precisión	Recall	Especificidad
Support Vector Machine	83.38%	99.05%	8.47%
Random Forest	84.94%	92.88%	46.9%
Artificial Neural Network	86.47%	92.91%	55.65%

Nota: Los datos fueron calculados usando el mismo conjunto de datos tanto para entrenamiento como para la prueba.

La tabla 6 indica los resultados se obtuvieron a través de la evaluación de tres modelos de clasificación, SVM, RF y ANN que fueron implementados para clasificar dengue.

En primer lugar, se comparó la precisión de los modelos, donde se observa que ANN presenta mayor desempeño con 86.47%, mostrando ser el más preciso para clasificar dengue, seguido del algoritmo RF tiene una precisión de 84.94% y el algoritmo SVM presenta

la menor precisión de 83.38%.

Seguido de ello, se comparó el recall de los modelos, siendo SVM el que mejor desempeño con 99.05% de recall, demostrando que modelo tiene la capacidad de identificar y recuperar un alto porcentaje de casos positivos, seguido de ANN que obtuvo recall de 92.91% y RF presenta menor recall de 92.88%.

Finalmente, se comparó la especificidad de los modelos, siendo ANN el que mejor desempeño con 55.65%, demostrando que modelo tiene la capacidad de identificar y clasificar correctamente un alto porcentaje de casos negativos en relación con el total de casos negativos, seguido de RF que obtuvo especificidad de 46.9% y el algoritmo SVM presenta menor recall de 8.47%

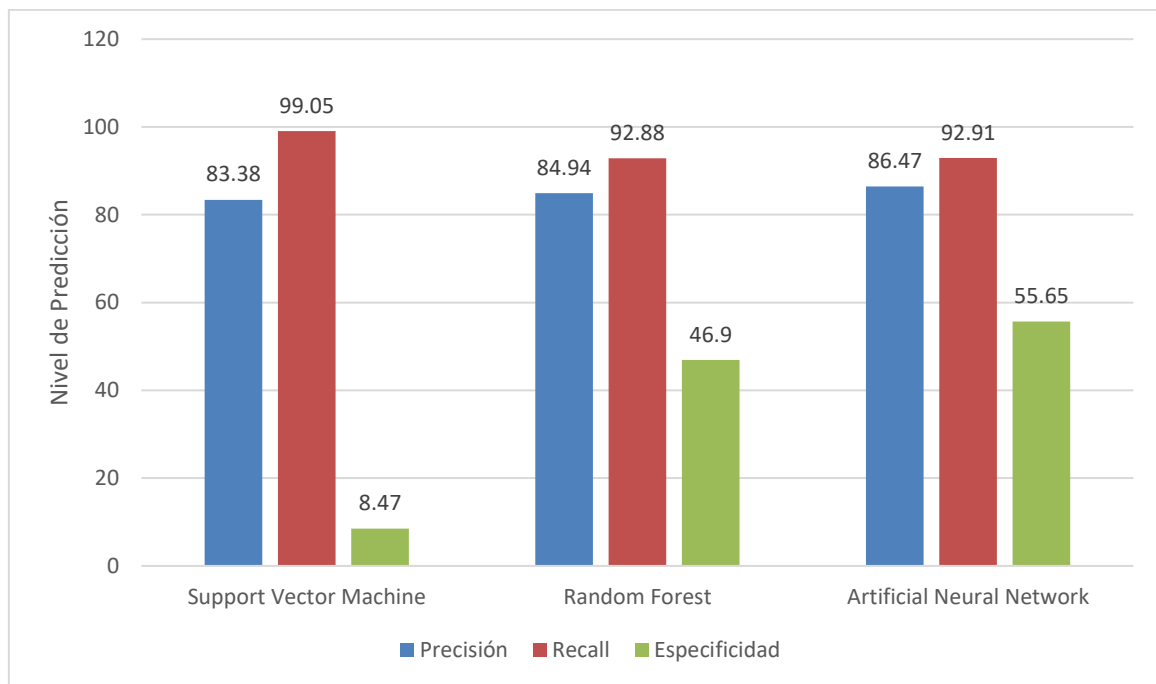


Fig. 17. Evaluación de rendimiento de los algoritmos de aprendizaje automático planteados en nuestra investigación

En la elección del algoritmo de aprendizaje automático con el mejor desempeño en la clasificación del dengue, se deben considerar tanto la precisión y eficacia en la identificación de casos positivos y negativos, así como la optimización de los recursos en el sector salud. Un algoritmo que sea eficiente y preciso no solo aumenta la detección de casos de dengue, sino que también optimiza los recursos y reduce los costos asociados con diagnósticos incorrectos o tratamientos innecesarios. Por lo tanto, se persigue un equilibrio entre

rendimiento y eficiencia para asegurar una gestión efectiva y sostenible de los recursos de salud.

3.2. Discusión

Los resultados alcanzados en la investigación, está enfocada en llevar a cabo un análisis comparativo de algoritmos de aprendizaje automático para la clasificación de dengue, brindan visiones importantes que contribuyen al avance de herramientas diagnósticas en la práctica clínica. La utilización de distintos algoritmos de aprendizaje automático mostraron resultados alentadores en cuanto a la precisión y recall en la clasificación de dengue.

El análisis llevado a cabo reveló que el algoritmo de aprendizaje automático SVM demostró una sensibilidad del 99.05%, cifra superior en comparación con otros enfoques. Esta característica facilita la optimización de los recursos al reconocer los algoritmos más eficientes, contribuye al reconocimiento de características y factores de riesgo relacionados con el dengue, ayudando así a la toma de decisiones clínicas. Además, aporta al progreso científico en el ámbito de la medicina.

No obstante, es importante abordar las instancias identificadas durante el estudio. La dataset presenta limitaciones en la generalidad de la información sobre síntomas, la categorización binaria de resultados que simplifica la complejidad clínica, y la agrupación de edades que podría no capturar la variabilidad intragrupal. La distinción autóctono o importado podría ser más específica, y la categorización mensual podría carecer de suficiente detalle temporal. La calidad y consistencia de los datos, posibles sesgos y consideraciones éticas son aspectos a abordar para mejorar la confiabilidad y utilidad de la dataset.

Comparando los resultados obtenidos con investigaciones previas, se observa que Mello-Roman et al. reportó un recall del 93% para el algoritmo SVM, cifra inferior a la obtenida en nuestra investigación. En su estudio, utilizaron una dataset de pacientes reales con diagnóstico previo de dengue, extraído del sistema de salud pública de Paraguay durante el período 2012-2016. Inicialmente, el conjunto de datos constaba de 4332 registros, pero tras el procesamiento quedaron 668 datos. Del total, el 90% destinaron para el entrenamiento y

el 10% para la prueba. Se llevaron a cabo diversas pruebas para determinar los mejores parámetros de las funciones de kernel y los parámetros óptimos de penalización, basados en la precisión de la clasificación en los conjuntos de datos de prueba. Por otro lado, Ozer et al. informó que el algoritmo SVM presentó un recall más bajo, alcanzando el 76%. En su estudio, utilizaron los conjuntos de datos SISA y SISAL, que incluyen información demográfica, historias médicas pasadas y datos de síntomas. La data SISA consta de 543 sujetos, con 28 características estimadas para cada registro, mientras que la data SISAL incluye 98 registros. Ambos conjuntos se dividieron para la validación tradicional en proporciones del 70% (entrenamiento) - 30% (prueba) y 85% (entrenamiento) - 15% (prueba), respectivamente, con el objetivo de predecir la necesidad de hospitalización de un paciente.

Las implicaciones prácticas derivadas de los resultados señalan que la aplicación de algoritmos de aprendizaje automático tiene el potencial de mejorar significativamente la clasificación del dengue. Sin embargo, la transición desde el ámbito de investigación hacia la aplicación clínica plantea desafíos adicionales. Estos desafíos incluyen la necesidad de integrarse con sistemas de información médica existentes y validar los resultados en contextos clínicos del mundo real.

El análisis proporciona información significativa acerca de la capacidad de los algoritmos de aprendizaje automático en la clasificación de dengue, resaltando la importancia de futuras investigaciones que aborden desafíos específicos o respalden la aplicación clínica de estas prometedoras herramientas.

IV. CONCLUSIONES Y RECOMENDACIONES

4.1. Conclusiones

- La construcción del conjunto de datos, enfocado en pacientes con síntomas específicos de Dengue, representa un recurso valioso para la investigación. Al identificar y recopilar datos detallados sobre los síntomas, se estableció una base robusta para analizar y comparar diferentes algoritmos de aprendizaje automático en la clasificación del Dengue.

- La revisión sistemática de la literatura permitió la elección de tres algoritmos destacados, evaluados con métricas de precisión. Dichos algoritmos se implementaron y evaluaron en un conjunto de entrenamiento (70%) y pruebas (30%), asegurando resultados confiables y robustos. El proceso metodológico respalda la validez de los hallazgos, destacando la aplicabilidad práctica de los algoritmos en la clasificación precisa del Dengue.

- En la evaluación de algoritmos de aprendizaje automático para clasificar el Dengue Virus, ANN sobresale como el líder, mostrando una impresionante precisión del 86.47%, un sólido recall del 92.91% y una moderada especificidad del 55.65%. En comparación, SVM destaca en la identificación de casos positivos pero tiene baja especificidad, mientras que RF ofrece un equilibrio sólido entre precisión y recall.

- Artificial Neural Network destaca como el algoritmo más efectivo en la clasificación del dengue, priorizando la eficiencia en el uso de recursos. Gracias a su alta precisión y especificidad, este método logra un equilibrio óptimo entre la identificación precisa de casos positivos de dengue y la minimización de los falsos positivos. Este enfoque permite un manejo más eficiente de los recursos disponibles para el tratamiento y control de la enfermedad.

4.2. Recomendaciones

No obstante a los alentadores resultados alcanzados con los algoritmos evaluados en este estudio, se recomienda la exploración y evaluación de otros algoritmos de aprendizaje automático para la clasificación del Dengue en futuras investigaciones. La diversidad de enfoques algorítmicos podría ofrecer nuevas perspectivas y mejoras en términos de precisión, recall y especificidad. Algoritmos como K-Nearest Neighbors y Logistic Regression, esos podrían aportar conocimientos valiosos y complementar los resultados obtenidos en este estudio. Se recomienda llevar a cabo comparaciones exhaustivas entre diferentes algoritmos para determinar cuáles ofrecen el mejor desempeño para clasificar dengue virus.

REFERENCIAS

- [1] T. Stephen and A. Rothman, "Infección por el virus del dengue: prevención y tratamiento," *UpToDate - Wolters Kluwer*, 2022.
- [2] S. Mattar, J. Montero A, and M. González Tous, "La historia del dengue aún no termina," *Rev MVZ Cordoba*, vol. 24, no. 2, pp. 7177–7179, Mar. 2019, doi: 10.21897/rmvz.1597.
- [3] M. M. Rubiano Godoy and B. S. Jiménez Cendales, "Validez diagnóstica de las pruebas clínicas para el diagnóstico diferencial de Dengue en población pediátrica.," *Pediatría (Bucur)*, vol. 53, no. 4, 2021, doi: 10.14295/rp.v53i4.166.
- [4] C. Y. Santos *et al.*, "A machine learning model to assess potential misdiagnosed dengue hospitalization," *Heliyon*, vol. 9, no. 6, p. e16634, Jun. 2023, doi: 10.1016/J.HELIYON.2023.E16634.
- [5] M. S. G. Shaikh, D. B. SureshKumar, and D. G. Narang, "Development of optimized ensemble classifier for dengue fever prediction and recommendation system," *Biomed Signal Process Control*, vol. 85, 2023, doi: 10.1016/j.bspc.2023.104809.
- [6] J. Yang, A. Al Mosabbir, E. Raheem, W. Hu, and M. S. Hossain, "Demographic characteristics, clinical symptoms, biochemical markers and probability of occurrence of severe dengue: A multicenter hospital-based study in Bangladesh," *PLoS Negl Trop Dis*, vol. 17, no. 3, 2023, doi: 10.1371/journal.pntd.0011161.
- [7] B. D. Handari, I. M. S. Niman, A. Hasan, J. R. P. Purba, and G. F. Hertono, "Comparation of elman neural network, long short-term memory, and gated recurrent unit in predicting dengue hemorrhagic fever at dki Jakarta," *Communications in Mathematical Biology and Neuroscience*, vol. 2021, 2021, doi: 10.28919/cmbn/6683.
- [8] F. Yavari Nejad and K. D. Varathan, "Identification of significant climatic risk factors and machine learning models in dengue outbreak prediction," *BMC Med Inform Decis Mak*, vol. 21, no. 1, 2021, doi: 10.1186/s12911-021-01493-y.
- [9] J. T. Lim, B. S. Dickens, S. Haoyang, N. L. Ching, and A. R. Cook, "Inference on dengue epidemics with Bayesian regime switching models," *PLoS Comput Biol*, vol. 16, no. 5,

- 2020, doi: 10.1371/journal.pcbi.1007839.
- [10] Z. Li, H. Gurgel, L. Xu, L. Yang, and J. Dong, "Improving Dengue Forecasts by Using Geospatial Big Data Analysis in Google Earth Engine and the Historical Dengue Information-Aided Long Short Term Memory Modeling," *Biology (Basel)*, vol. 11, no. 2, 2022, doi: 10.3390/biology11020169.
- [11] I. N. Tanawi, V. Vito, D. Sarwinda, H. Tasman, and G. F. Hertono, "Support Vector Regression for Predicting the Number of Dengue Incidents in DKI Jakarta," in *Procedia Computer Science*, 2021. doi: 10.1016/j.procs.2021.01.063.
- [12] S. Kaur and S. Sharma, "Comparative Analysis of Machine Learning Classifiers on Forecasting Dengue Fever Infection," in *Advances in Transdisciplinary Engineering*, 2023. doi: 10.3233/ATDE221302.
- [13] H. Hamdani, H. R. Hatta, N. Puspitasari, A. Septiarini, and Henderi, "Dengue classification method using support vector machines and cross-validation techniques," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 3, 2022, doi: 10.11591/ijai.v11.i3.pp1119-1129.
- [14] L. M. Stolerman, P. D. Maia, and J. Nathan Kutz, "Forecasting dengue fever in Brazil: An assessment of climate conditions," *PLoS One*, vol. 14, no. 8, 2019, doi: 10.1371/journal.pone.0220106.
- [15] Suhaeri, N. M. Nawi, and M. Fathurahman, "Early detection of Dengue disease using extreme learning machine," *Int J Adv Sci Eng Inf Technol*, vol. 8, no. 5, 2018, doi: 10.18517/ijaseit.8.5.5006.
- [16] R. Kapoor, S. Ahuja, and V. Kadyan, "Machine Learning Based Classification Algorithm for Classification of Dengue (Dengue Fever - DF, Dengue Harmonic Fever - DHF, Serbe Dengue - SD)," *ECS Trans*, vol. 107, no. 1, 2022, doi: 10.1149/10701.4659ecst.
- [17] M. Saleem, S. Ali, M. Bilal, K. Safdar, and M. Hassan, "Development of multivariate classification models for the diagnosis of dengue virus infection," *Photodiagnosis Photodyn Ther*, vol. 40, 2022, doi: 10.1016/j.pdpdt.2022.103136.
- [18] O. Mudele, A. Frery, L. Zanandrez, A. Eiras, and P. Gamba, "Dengue Vector Population

- Forecasting Using Multisource Earth Observation Products and Recurrent Neural Networks,” *IEEE J Sel Top Appl Earth Obs Remote Sens*, vol. 14, 2021, doi: 10.1109/JSTARS.2021.3073351.
- [19] A. Rahman and S. S. Prasetyowati, “Performance Analysis of the Hybrid Voting Method on the Classification of the Number of Cases of Dengue Fever,” *International Journal on Information and Communication Technology (IJoICT)*, vol. 8, no. 1, 2022, doi: 10.21108/ijoict.v8i1.614.
- [20] S. Shenoy *et al.*, “Artificial intelligence in differentiating tropical infections: A step ahead,” *PLoS Negl Trop Dis*, vol. 16, no. 6, 2022, doi: 10.1371/JOURNAL.PNTD.0010455.
- [21] G. M. E. Hair, F. F. Nobre, and P. Brasil, “Characterization of clinical patterns of dengue patients using an unsupervised machine learning approach,” *BMC Infect Dis*, vol. 19, no. 1, 2019, doi: 10.1186/s12879-019-4282-y.
- [22] R. Aguas, I. Dorigatti, L. Coudeville, C. Luxemburger, and N. M. Ferguson, “Cross-serotype interactions and disease outcome prediction of dengue infections in Vietnam,” *Sci Rep*, vol. 9, no. 1, 2019, doi: 10.1038/s41598-019-45816-6.
- [23] M. S. Rahman *et al.*, “Mapping the spatial distribution of the dengue vector *Aedes aegypti* and predicting its abundance in northeastern Thailand using machine-learning approach,” *One Health*, vol. 13, 2021, doi: 10.1016/j.onehlt.2021.100358.
- [24] S. Ismail, R. Fildes, R. Ahmad, W. N. Wan Mohamad Ali, and T. Omar, “The practicality of Malaysia dengue outbreak forecasting model as an early warning system,” *Infect Dis Model*, vol. 7, no. 3, 2022, doi: 10.1016/j.idm.2022.07.008.
- [25] I. Sanchez-Gendriz *et al.*, “Data-driven computational intelligence applied to dengue outbreak forecasting: a case study at the scale of the city of Natal, RN-Brazil,” *Sci Rep*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-10512-5.
- [26] W. Hoyos, J. Aguilar, and M. Toro, “An autonomous cycle of data analysis tasks for the clinical management of dengue,” *Heliyon*, vol. 8, no. 10, 2022, doi: 10.1016/j.heliyon.2022.e10846.

- [27] R. Gangula, L. Thirupathi, R. Parupati, K. Sreeveda, and S. Gattoju, "Ensemble machine learning based prediction of dengue disease with performance and accuracy elevation patterns," *Mater Today Proc*, 2021, doi: 10.1016/j.matpr.2021.07.270.
- [28] I. Ozer, O. Cetin, K. Gorur, and F. Temurtas, "Improved machine learning performances with transfer learning to predicting need for hospitalization in arboviral infections against the small dataset," *Neural Comput Appl*, vol. 33, no. 21, 2021, doi: 10.1007/s00521-021-06133-0.
- [29] J. D. Mello-Román, J. C. Mello-Román, S. Gómez-Guerrero, and M. García-Torres, "Predictive Models for the Medical Diagnosis of Dengue: A Case Study in Paraguay," *Comput Math Methods Med*, vol. 2019, 2019, doi: 10.1155/2019/7307803.
- [30] F. Rustam *et al.*, "Vector mosquito image classification using novel RIFS feature selection and machine learning models for disease epidemiology," *Saudi J Biol Sci*, vol. 29, no. 1, 2022, doi: 10.1016/j.sjbs.2021.09.021.
- [31] T. T. Han *et al.*, "Machine learning based classification model for screening of infected patients using vital signs," *Inform Med Unlocked*, vol. 24, 2021, doi: 10.1016/j.imu.2021.100592.
- [32] R. Zargari Marandi *et al.*, "Development of a machine learning model for early prediction of plasma leakage in suspected dengue patients," *PLoS Negl Trop Dis*, vol. 17, no. 3, 2023, doi: 10.1371/journal.pntd.0010758.
- [33] S. Choubey, S. Barde, and A. Badholia, "Enhancing the prediction efficiency of virus borne diseases using enhanced backpropagation with an artificial neural network," *Measurement: Sensors*, vol. 24, 2022, doi: 10.1016/j.measen.2022.100505.
- [34] W. Nadda, W. Boonchieng, and E. Boonchieng, "Influenza, dengue and common cold detection using LSTM with fully connected neural network and keywords selection," *BioData Min*, vol. 15, no. 1, 2022, doi: 10.1186/s13040-022-00288-9.
- [35] X. Zhao, K. Li, C. K. E. Ang, and K. H. Cheong, "A deep learning based hybrid architecture for weekly dengue incidences forecasting," *Chaos Solitons Fractals*, vol. 168, 2023, doi: 10.1016/j.chaos.2023.113170.

- [36] M. S. Fernandes, W. Cordeiro, and M. Recamonde-Mendoza, "Detecting *Aedes aegypti* mosquitoes through audio classification with convolutional neural networks," *Comput Biol Med*, vol. 129, 2021, doi: 10.1016/j.combiomed.2020.104152.
- [37] D. J. Matuszewski and I. M. Sintorn, "TEM virus images: Benchmark dataset and deep learning classification," *Comput Methods Programs Biomed*, vol. 209, 2021, doi: 10.1016/j.cmpb.2021.106318.
- [38] Y. Sukama, G. F. Hertono, B. D. Handari, and D. Aldila, "Comparing activation functions in predicting dengue hemorrhagic fever cases in DKI Jakarta using u," in *AIP Conference Proceedings*, 2020. doi: 10.1063/5.0030456.
- [39] M. C. D. Santos *et al.*, "Infrared spectroscopy (NIRS and ATR-FTIR) together with multivariate classification for non-destructive differentiation between female mosquitoes of *Aedes aegypti* recently infected with dengue vs. uninfected females," *Acta Trop*, vol. 235, 2022, doi: 10.1016/j.actatropica.2022.106633.

ANEXOS:

Instrumento de recolección de datos: Ficha electrónica de Observación

Total de casos:

Confirmados	17491
Descartados	3666
TOTAL	21157

Confirmados:

Sexo	N° de Casos
Femenino	9536
Masculino	7955
Total	17491

Edades	N° Casos
Mayor 60 años	1833
30 a 59 años - ADULTO	4750
18 a 29 años - JOVEN	1149
12 a 17 años - ADOLESCENTE	779
00 a 11 años - NIÑO	1025
Total	9536

Edades	N° Casos
Mayor 60 años	1476
30 a 59 años - ADULTO	3522
18 a 29 años - JOVEN	1086
12 a 17 años - ADOLESCENTE	821
00 a 11 años - NIÑO	1050
Total	7955

Origen	N° de Casos
Autóctonos	14846
Importados	2645
Total	17491

Autoctono	N° Casos
Chiclayo	5160
Ferreñafe	2978
José Leonardo Ortiz	1459
Tumán	1023
Cayalti	902
La Victoria	595
Jayanca	594
Lambayeque	416
Oyotun	316
Motupe	293
Eten	286
Chongoyape	264
Olmos	145
Pomalca	121
Pátapo	56
Pimentel	34
Reque	28
Pucalá	27
Picsi	18
Monsefú	16
OTROS	115
TOTAL	14846

Importado	N° Casos
Chepén	2398
Guadalupe	77
Jaén	41
Chota	22
Bagua	18
Lima	13
Piura	12
Cajamarca	9
Chiclayo	5
Trujillo	5
Utcubamba	3
Santa Cruz	3
San Ignacio	3
Sullana	2
Otros	34
TOTAL	2645

Descartados:

Sexo	N° de Casos
Femenino	1861
Masculino	1985
Total	3666

Edades	N° Casos
Mayor 60 años	82
30 a 59 años - ADULTO	459
18 a 29 años - JOVEN	379
12 a 17 años - ADOLESCENTE	396
00 a 11 años - NIÑO	365
Total	1861

Edades	N° Casos
Mayor 60 años	92
30 a 59 años - ADULTO	553
18 a 29 años - JOVEN	480
12 a 17 años - ADOLESCENTE	442
00 a 11 años - NIÑO	418
Total	1985

Origen	N° de Casos
Autóctonos	1735
Importados	1931
Total	3666

Instrumento de Validación



Universidad
Señor de Sipán

INSTRUMENTO PARA VALIDACIÓN: ENCUESTA

TEMA: Análisis Comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus

Marca con **X** la respuesta que Usted crea conveniente

Ítem	Pregunta	Si	No
1	¿Considera que los grupos etarios utilizados en la investigación son representativos y adecuados para la investigación?	X	
2	¿Cree usted que los síntomas que se han considerado en la investigación son los adecuados?	X	
3	¿Considera que los resultados obtenidos tienen relevancia práctica en el campo de la epidemiología del dengue?	X	
4	¿Cree que los resultados obtenidos pueden ser útiles en la práctica clínica para ayudar en el diagnóstico y tratamiento del dengue?	X	
5	¿Considera que los resultados obtenidos en la investigación podrían influir en su práctica clínica diaria relacionada con el manejo del dengue?	X	
6	¿Confía en los resultados presentados por cada algoritmo para ayudar en la detección y clasificación del virus del dengue?	X	

Muchas Gracias por su respuesta



Dr. Jorge Fernández Mogollón
MÉDICO ASISTENTE C.M.P. 5.194
OFICINA DE INTELIGENCIA SANITARIA
EsSalud RED PRESTACIONAL
LAMBAEQUE

FERNANDEZ MOGOLLON JORGE LUIS

MEDICO CIRUJANO



Identificadores

Orcid Id [0000-0001-8293-0882](#)
Scopus Id [57192830531](#)
Wos Id

Renacyt

Código de Registro P0055907
Nivel V
Fecha de Registro 27/09/2023
Estado Activo al 18/01/2024
Descargar Constancia [Constancia](#)

Contacto

Teléfono
Correo jorgeluis_59@yahoo.es



Fecha de última actualización: 21/08/2023

Ficha CTI Vitae: [55907](#)

Fecha de exportación: 18/01/2024 09:25:36

Resumen

Médico cirujano. Asistente de la Oficina de Inteligencia Sanitaria. Red Prestacional Lambayeque- EsSalud. Egresado de la Maestría en Epidemiología Clínica de la Universidad Peruana Cayetano Heredia. Docente de la Facultad de Medicina de la Universidad Señor de Sipán. Anteriormente se ha desempeñado como médico encargado del trabajo de campo para control del Covid-19 en la Gerencia Regional de Salud Lambayeque. Ha realizado docencia en investigación en las facultades de Medicina de la Universidad de San Martín de Porres. Filial norte, Universidad Santo Toribio de Mogrovejo USAT y Universidad de Chiclayo. Es miembro revisor de la revista Experiencia en Medicina (Hospital Regional Lambayeque), Revista del Cuerpo Médico del HNAAA y En la Revista Panamericana de Salud Pública (OPS/OMS).

Datos Personales

Sexo	Masculino
Tipo de Documento de Identidad	DNI
Número de Documento de Identidad	43284780
País de Nacimiento	Perú
Fecha de Nacimiento	05/09/1985
Dirección	Psj.San Sebastian 189
Departamento	LAMBAYEQUE
Provincia	LAMBAYEQUE
Distrito	LAMBAYEQUE

Datos Actuales

Página web personal	http://
E-mail	jorgeluis_59@yahoo.es
Dirección actual	Psj.San Sebastian 189
País	Perú
Teléfono de contacto	
Celular	
Departamento	Lambayeque
Provincia	Lambayeque
Distrito	Lambayeque

Otros Identificadores

Scopus Author ID	57192830531
ORCID ID	0000-0001-8293-0882
Web of Science	

Experiencias

Cargo	Cargo en I+D+i	Institución	Fecha de Inicio	Fecha de fin	Trabajo actualment e	Descripción del cargo
Medico	Otros cargos relacionados a (I+D+i)	Gerencia Regional de Salud de Lambayeque	2021	2022	No	Médico encargado de las Unidades Móviles de Atención Especializada a Población Excluida (UMAEP)
Médico Epidemiólogo	Otros cargos relacionados a (I+D+i)	Hospital Almanzor Aguinaga Asenjo	2013	2021	No	Oficina inteligencia sanitaria
Médico	Otros cargos relacionados a (I+D+i)	Clinica del Pacifico S.A.	2012	2014	No	Médico área de emergencias. Médico auditor
Medico Epidemiólogo	Otros cargos relacionados a (I+D+i)	Hospital Almanzor Aguinaga Asenjo	2023	Presente	Si	Médico Epidemiólogo de la Oficina de Inteligencia Sanitaria

Experiencia como docente

Institución	Tipo de Institución	Tipo de docente universitario	Descripción del cargo	Fecha de Inicio	Fecha de fin	Trabajo actualment e
Universidad Católica Santo Toribio de Mogrovejo	Universidad	Contratado	Docente Contratado	2021	Presente	Si
Universidad Señor de Sipan S.A.C.	Universidad	Contratado	Docente Contratado	2022	Presente	Si
Universidad de San Martín de Porres	Universidad	Contratado	Docente Contratado	2015	2019	No
Universidad Particular de Chiclayo	Universidad	Contratado	Docente Contratado	2016	2017	No
Universidad Señor de Sipan Sac	Universidad	Contratado	Docente Contratado	2012	2013	No

Experiencia como asesor de tesis

Universidad	Tesis	Tesista	Fecha de aceptación de tesis	Enlace de Fuente del repositorio académico
Universidad Santo Toibio de Mogrovejo	Licenciado / Título	César Zapata Zapata / Francisco Fernández Rioja	Marzo 2015	https://tesis.usat.edu.pe/handle/20.500.12423/306
Universidad Católica Santo Toribio de Mogrovejo	Licenciado / Título	Neciosup Lopez Giancarlo	Mayo 2022	https://tesis.usat.edu.pe/handle/20.500.12423/4632
Universidad de San Martín de Porres	Licenciado / Título	ROJAS AQUINO, RUBY	Febrero 2017	http://www.repositorioacademico.usmp.edu.pe/handle/usmp/4428
Universidad de San Martín de Porres	Licenciado / Título	RODRIGO RIVERA, CRISTIAN	Febrero 2017	http://www.repositorioacademico.usmp.edu.pe/handle/usmp/4430
Universidad de San Martín de Porres	Licenciado / Título	VERONA BALCAZAR, SEGUNDO MAURICIO	Febrero 2017	http://www.repositorioacademico.usmp.edu.pe/handle/usmp/



Universidad
Señor de Sipán

INSTRUMENTO PARA VALIDACIÓN: ENCUESTA

TEMA: Análisis Comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus

Marca con **X** la respuesta que Usted crea conveniente

Ítem	Pregunta	Si	No
1	¿Considera que los grupos etarios utilizados en la investigación son representativos y adecuados para la investigación?	X	
2	¿Cree usted que los síntomas que se han considerado en la investigación son los adecuados?	X	
3	¿Considera que los resultados obtenidos tienen relevancia práctica en el campo de la epidemiología del dengue?	X	
4	¿Cree que los resultados obtenidos pueden ser útiles en la práctica clínica para ayudar en el diagnóstico y tratamiento del dengue?	X	
5	¿Considera que los resultados obtenidos en la investigación podrían influir en su práctica clínica diaria relacionada con el manejo del dengue?	X	
6	¿Confía en los resultados presentados por cada algoritmo para ayudar en la detección y clasificación del virus del dengue?	X	

Muchas Gracias por su respuesta


Msc. Lic. Elizabeth V. Neciosup Puican
CEP 23102 RE 8341 RM 273
JEFE OFICINA DE INTELIGENCIA SANITARIA
RED PRESTACIONAL LAMSAVEQUE
Firma



23102

Buscar

Apellidos y nombres

Buscar



NECIOSUP PUICAN ELIZABETH VICTORIA

CONSEJO REGIONAL IX LAMBAYEQUE

✓ HABIL

#	Grado	#Registro	Nombre
1	Especialidad	008341	ENFERMERIA EN SALUD PUBLICA Y COMUNITARIA
2	Maestría	000273	CIENCIAS CON MENCION EN SALUD PUBLICA
3	Especialidad	010702	EPIDEMIOLOGIA



Universidad
Señor de Sipán

INSTRUMENTO PARA VALIDACIÓN: ENCUESTA

TEMA: Análisis Comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus

Marca con **X** la respuesta que Usted crea conveniente

Ítem	Pregunta	Si	No
1	¿Considera que los grupos etarios utilizados en la investigación son representativos y adecuados para la investigación?	X	
2	¿Cree usted que los síntomas que se han considerado en la investigación son los adecuados?	X	
3	¿Considera que los resultados obtenidos tienen relevancia práctica en el campo de la epidemiología del dengue?	X	
4	¿Cree que los resultados obtenidos pueden ser útiles en la práctica clínica para ayudar en el diagnóstico y tratamiento del dengue?	X	
5	¿Considera que los resultados obtenidos en la investigación podrían influir en su práctica clínica diaria relacionada con el manejo del dengue?	X	
6	¿Confía en los resultados presentados por cada algoritmo para ayudar en la detección y clasificación del virus del dengue?	X	

Muchas Gracias por su respuesta

GOBIERNO REGIONAL LAMBAYEQUE
GERENCIA REGIONAL DE SALUD


Dr. Eduardo Fidel Vergara Wekselman
JEFE DE OFICINA DE EPIDEMIOLOGIA
Firma



CMP	Apellidos	Nombres
017171	VERGARA WEKSELMAN	EDUARDO FIDEL

HÁBIL

Foto	Email	Consejo Regional
		CONSEJO REGIONAL VIII CHICLAYO

Registro	Tipo	Código	Fecha
EPIDEMIOLOGIA DE CAMPO	RNE	020215	

Buscar Otro



Consentimiento informado:



Comité Institucional De Ética En Investigación (CIEI-RPLAMB)

NOTA N° 024 - CIEI-RPLAMB-ESSALUD-2024
Chiclayo, 23 de febrero del 2024

Señorita
Exebio Chepe Yolanda Victoria de Fátima.
Presente. -

ASUNTO: Aprobación de proyecto de investigación: "Análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus"

- La presente es para saludarla cordialmente así mismo informarle que, el proyecto de investigación señalado a continuación fue **APROBADO**:
 - **Título del Estudio:** "Análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus"
 - **Investigador Principal(es):** Exebio Chepe Yolanda Victoria de Fatima
 - **Asesor (es Mg. Bravo Ruiz Jaime Arturo**
 - **Colaborador:** Jorge Luis Fernández Mogollón
 - **Institución:** UNIVERSIDAD SEÑOR DE SIPAN. FACULTAD DE INGENIERIA, ARQUITECTURA Y URBANISMO. ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS
 - **Condición:** Aprobado
 - **Periodo de permiso de ejecución:** 23 de febrero al 23 de julio del 2024

Los documentos que se revisaron y aprobaron fueron;

- Protocolo de investigación: "Análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus"
- Solicitud para evaluación de protocolo de investigación. Anexo A
- Carta de aceptación para la realización de la investigación por el Jefe del Departamento / Servicio / Área o Jefe inmediato superior. (Anexo B)
- Declaración Jurada (indicando que el desarrollo del estudio de investigación no irrogará gastos a EsSalud). Anexo C.
- Boleta electrónica N° B145-00004014
- Carta de Compromiso del Investigador Principal. (Anexo E)


Dr. Stalin Tello Vera
PRESIDENTE TITULAR
COMITÉ INSTITUCIONAL DE ÉTICA EN INVESTIGACIÓN
RED PRESTACIONAL LAMBAYEQUE


Carta de autorización para la recolección de la información:



RED PRESTACIONAL LAMBAYEQUE
HOSPITAL NACIONAL ALMANZOR AGUINAGA ASENJO

Comité Institucional De Ética En Investigación (CIEI-RPLAMB)

CERTIFICADO DE APROBACIÓN ÉTICA: N° 007- CIEI-RPLAMB.2024

El Presidente del Comité Institucional de Ética en Investigación de la Red Prestacional Lambayeque, hace constar que, el protocolo de investigación señalado a continuación fue **APROBADO**:

- **Título del Estudio:** "Análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus"
- **Investigador Principal(es):** Exebio Chepe Yolanda Victoria de Fatima
- **Asesor (es Mg. Bravo Ruiz Jaime Arturo**
- **Colaborador:** Jorge Luis Fernández Mogollón
- **Institución:** UNIVERSIDAD SEÑOR DE SIPAN. FACULTAD DE INGENIERIA, ARQUITECTURA Y URBANISMO. ESCUELA PROFESIONAL DE INGENIERIA DE SISTEMAS
- **Condición:** Aprobado
- **Periodo de permiso de ejecución:** 16 de febrero al 16 de junio del 2024

Para la aprobación se ha considerado el cumplimiento de pautas éticas en investigación, incluyendo el balance beneficio/riesgo, confidencialidad de los datos y otros.

Cualquier enmienda en los objetivos secundarios, metodología y aspectos éticos debe ser solicitada a este comité.

El periodo de vigencia de la presente aprobación será de **05 meses**; desde el 23 de febrero al 23 de julio del 2024, debiendo solicitar la renovación con 30 días de anticipación.

Sírvase enviar el informe final del estudio, una vez concluido el mismo a la Oficina de Investigación y Docencia del Hospital Nacional Almanzor Aguinaga Asenjo.

Chiclayo, 23 de febrero del 2024

STV/cva
(CIEI-RPLAMB.: 23-FEB-2024)
Folios: ()

NIT	7550	2024	010
-----	------	------	-----

Dr. Stalin Tello Vera
PRESIDENTE TITULAR
COMITÉ INSTITUCIONAL DE ÉTICA EN INVESTIGACIÓN
RED PRESTACIONAL LAMBAYEQUE
EsSalud

Evidencias de ejecución:

- **Algoritmo: Support Vector Machine**

```
info_memoria = psutil.virtual_memory()

# Mostrar información
print(f"Memoria total: {info_memoria.total / (1024 ** 3):.2f} GB")
print(f"Memoria disponible: {info_memoria.available / (1024 ** 3):.2f} GB")
print(f"Memoria utilizada: {info_memoria.used / (1024 ** 3):.2f} GB")
print(f"Porcentaje de memoria utilizada: {info_memoria.percent}%")

print('Promedio de tiempo de respuesta: {0:5f}seg:'.format(fin -start))
print('Grado de consumo de CPU:',psutil.cpu_percent(interval=None), "%")
```

Precisión: 83.38 %
Recall: 99.05 %
Especificidad: 8.47 %
Memoria total: 12.68 GB
Memoria disponible: 11.35 GB
Memoria utilizada: 1.06 GB
Porcentaje de memoria utilizada: 10.5%
Promedio de tiempo de respuesta: 0.011517seg:
Grado de consumo de CPU: 35.4 %

- **Algoritmo: Random Forest**

```
# Mostrar información
print(f"Memoria total: {info_memoria.total / (1024 ** 3):.2f} GB")
print(f"Memoria disponible: {info_memoria.available / (1024 ** 3):.2f} GB")
print(f"Memoria utilizada: {info_memoria.used / (1024 ** 3):.2f} GB")
print(f"Porcentaje de memoria utilizada: {info_memoria.percent}%")

print('Promedio de tiempo de respuesta: {0:5f}seg:'.format(fin -start))
print('Grado de consumo de CPU:',psutil.cpu_percent(interval=None), "%")
```

⇒ Precisión: 84.94 %
Recall: 92.88 %
Especificidad: 46.99 %
Memoria total: 12.68 GB
Memoria disponible: 11.41 GB
Memoria utilizada: 1.00 GB
Porcentaje de memoria utilizada: 10.0%
Promedio de tiempo de respuesta: 0.024931seg:
Grado de consumo de CPU: 21.0 %

- **Algoritmo: Artificial Neural Network**

```
[ ] # Mostrar información
print(f"Memoria total: {info_memoria.total / (1024 ** 3):.2f} GB")
print(f"Memoria disponible: {info_memoria.available / (1024 ** 3):.2f} GB")
print(f"Memoria utilizada: {info_memoria.used / (1024 ** 3):.2f} GB")
print(f"Porcentaje de memoria utilizada: {info_memoria.percent}%")

print('Promedio de tiempo de respuesta: {0:5f}seg:'.format(fin -start))
print('Grado de consumo de CPU:',psutil.cpu_percent(interval=None), "%")
```

```
Precisión: 86.47 %
Recall: 92.91 %
Especificidad: 55.65 %
Memoria total: 12.68 GB
Memoria disponible: 11.09 GB
Memoria utilizada: 1.32 GB
Porcentaje de memoria utilizada: 12.5%
Promedio de tiempo de respuesta: 0.051338seg:
Grado de consumo de CPU: 62.3 %
```

Operacionalización de las variable

Variable de estudio	Definición conceptual	Definición operacional	Indicadores	Instrumento	Valores finales	Tipo de variable	Escala de medición
Algoritmos de aprendizaje automático.	Rama de la Inteligencia Artificial, crear y desarrollar algoritmos específicos que obtengan resultados de alto rendimiento.	<ol style="list-style-type: none"> Definición de variables Selección de métricas Recopilación de datos Preprocesamiento de datos Aplicar algoritmos 	<ul style="list-style-type: none"> Consumo de Memoria Tiempo de respuesta Consumo de CPU 	Ficha electrónica de observación	Segundo Gigabyte o byte	Independiente	El algoritmo K-NN obtuvo un rendimiento de memoria
Clasificar Dengue Virus	Enfermedad viral transmitida por el mosquito Aedes aegypti. Su transmisión es través de la picadura de mosquitos infectados	<ol style="list-style-type: none"> Recopilación de datos Preprocesamiento de datos Selección del algoritmo Entrenamiento del modelo Validación Interpretación de resultados Información de resultados 	<ul style="list-style-type: none"> Precisión Recall Especificidad 	Ficha electrónica de observación	Rendimiento (%) Margen de error	Dependiente	Escala ordinal

NOMBRE DEL TRABAJO

Análisis comparativo de algoritmos de aprendizaje automático para clasificar Dengue Virus

AUTOR

Yolanda Victoria de Fatima Exebio Chepe

RECUENTO DE PALABRAS

4717 Words

RECUENTO DE CARACTERES

26525 Characters

RECUENTO DE PÁGINAS

24 Pages

TAMAÑO DEL ARCHIVO

1.6MB

FECHA DE ENTREGA

Apr 29, 2024 2:58 PM GMT-5

FECHA DEL INFORME

Apr 29, 2024 2:59 PM GMT-5**● 12% de similitud general**

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos.

- 7% Base de datos de Internet
- Base de datos de Crossref
- 9% Base de datos de trabajos entregados
- 3% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Coincidencia baja (menos de 8 palabras)
- Material citado