



Unidad Pedagógica
San Ignacio Guazú

CARRERA DE INGENIERÍA INFORMÁTICA

TRABAJO FINAL DE GRADO

**SIMULACIÓN DINÁMICA EN TIEMPO REAL DEL COMPORTAMIENTO DE
LA PROPAGACIÓN DEL DENGUE A TRAVÉS DE MACHINE LEARNING**

2024



Unidad Pedagógica
San Ignacio Guazú

CARRERA DE INGENIERÍA INFORMÁTICA

TRABAJO FINAL DE GRADO

SIMULACIÓN DINÁMICA EN TIEMPO REAL DEL COMPORTAMIENTO DE LA PROPAGACIÓN DEL DENGUE A TRAVÉS DE MACHINE LEARNING

Fabián Rolón Pedrozo

Tutores:

Ing. Santiago Caballero Semenaka

Ing. Aura Rocío Villalba

2024

Dedicatoria

A Dios Todopoderoso, por darme la fortaleza y la determinación necesarias para seguir adelante en este camino.

A mis padres, quienes son los pilares fundamentales de mi vida, por su amor incondicional y los sacrificios que han hecho para que pueda formarme. Sus enseñanzas y su guía han sido invaluable en cada paso de este recorrido.

A mi hermano, ejemplo vivo de perseverancia y fortaleza, por su apoyo constante y por inspirarme a ser mejor cada día.

A mi novia, mi compañera incondicional en este viaje, por su amor, paciencia y comprensión. Su presencia ha sido crucial en los momentos más desafiantes de este proceso.

A mis profesores ingenieros, por compartir sus conocimientos y experiencias con dedicación y profesionalismo. Su guía ha sido esencial en mi formación académica y desarrollo profesional.

Agradecimientos

Quiero expresar mi más sincero agradecimiento, en primer lugar, a Dios por darme la fortaleza y la perseverancia necesarias en cada paso de este camino.

A mis padres, por su amor, paciencia y apoyo incondicional a lo largo de toda mi formación académica. Sin su constante aliento y sacrificios, este logro no habría sido posible. Son el pilar fundamental que me ha permitido llegar hasta aquí.

A mi hermano, por su invaluable compañerismo y por ser una fuente constante de inspiración y motivación. Su ejemplo ha sido clave en mi desarrollo personal y profesional.

A mi novia, compañera incondicional en esta travesía, por su amor y paciencia inquebrantables. Gracias por ser mi soporte emocional y por estar presente en cada momento, compartiendo tanto los desafíos como las alegrías de este proceso.

A mis profesores ingenieros, por su excepcional dedicación y compromiso con mi educación. Su valioso conocimiento, orientación en los momentos difíciles y constante confianza en mis capacidades han sido fundamentales para superar cada obstáculo en este camino.

Finalmente, extendiendo mi gratitud a todas aquellas personas que, de una u otra manera, contribuyeron a hacer de este proyecto una realidad. Su apoyo y colaboración han sido invaluable en la consecución de esta meta.

Contenido del libro

AGRADECIMIENTOS.....	II
CONTENIDO DEL LIBRO	III
ÍNDICE DE FIGURAS.....	V
ÍNDICE DE TABLAS.....	XI
RESUMEN	XII
ABSTRACT.....	XIII
1. INTRODUCCIÓN.....	1
1.1. PLANTEAMIENTO DEL PROBLEMA	4
1.2. JUSTIFICACIÓN.....	7
1.3. ALCANCE Y LIMITACIONES.....	8
1.4. OBJETIVOS.....	10
1.4.1. <i>Objetivo general</i>	10
1.4.2. <i>Objetivos específicos</i>	11
1.5. CONTENIDO DEL LIBRO	11
2. MARCO TEÓRICO.....	13
2.1. ANTECEDENTES DE INVESTIGACIÓN.....	13
2.2. BASES TEÓRICAS.....	20
2.2.1. <i>Antecedentes del dengue</i>	20
2.2.2. <i>Impacto del dengue en la salud pública del Paraguay</i>	21
2.2.3. <i>Modelos Matemáticos y Computacionales de propagación de enfermedades</i>	23
2.2.4. <i>Evaluación del rendimiento de simulaciones dinámicas</i>	27
2.2.5. <i>Algoritmos de ML</i>	28
2.2.6. <i>Lenguaje de programación Python</i>	29
2.2.7. <i>Marco Legal</i>	30

3.	MARCO METODOLÓGICO.....	34
3.1.	ENFOQUE DE INVESTIGACIÓN	34
3.2.	DISEÑO DE INVESTIGACIÓN	35
3.3.	TIPOS DE LA INVESTIGACIÓN.....	35
3.4.	ETAPAS METODOLÓGICAS.....	36
3.4.1.	<i>Etapa 1: Exploratoria.....</i>	<i>36</i>
3.4.2.	<i>Etapa 2: Diseño e Integración</i>	<i>55</i>
3.4.3.	<i>Etapa 3: Pruebas.....</i>	<i>73</i>
3.5.	ANÁLISIS DE RESULTADOS	84
3.5.1.	<i>Comparación de resultados entre predicción y datos reales.....</i>	<i>84</i>
3.5.2.	<i>Evaluación del rendimiento del sistema</i>	<i>84</i>
3.5.3.	<i>Red Neuronal LSTM (Long Short-Term Memory).....</i>	<i>85</i>
3.5.4.	<i>Pruebas</i>	<i>86</i>
3.5.5.	<i>Visualización de datos históricos separadas por año.....</i>	<i>96</i>
4.	CONCLUSIONES Y TRABAJOS FUTUROS.....	105
5.	ANEXOS.....	106
5.1.	CARTA DE AUTORIZACIÓN PARA LA INVESTIGACIÓN	106
5.2.	ENTREGA DE DATOS EPIDEMIOLÓGICOS PARA LA INVESTIGACIÓN	107
5.3.	APROBACIÓN DEL PROTOTIPO	108
5.4.	AVANCES DEL TRABAJO FINAL DE GRADO.....	109
6.	GLOSARIO DE TÉRMINOS Y ABREVIATURAS.....	114
7.	REFERENCIAS BIBLIOGRÁFICAS	117

Índice de Figuras

Figura Nro 1 <i>Recolección de datos epidemiológicos 2022- 2024</i>	37
Figura Nro 2 <i>Archivo de datos epidemiológicos diciembre 2023</i>	38
Figura Nro 3 <i>Planilla de datos meteorológicos marzo 2024</i>	38
Figura Nro 4 <i>Planilla de datos meteorológicos 2019 – 2023</i>	39
Figura Nro 5 <i>Planilla de datos meteorológicos abril 2024</i>	39
Figura Nro 6 <i>Datos meteorológicos – año 2022</i>	40
Figura Nro 7 <i>Datos cargados en Excel</i>	40
Figura Nro 8 <i>Datos cargados en Excel (continuación)</i>	41
Figura Nro 9 <i>Datos cargados en Excel (continuación)</i>	41
Figura Nro 10 <i>Datos epidemiológicos – abril 2024</i>	42
Figura Nro 11 <i>Datos epidemiológicos 2022</i>	42
Figura Nro 12 <i>Preparación de datos para cargar en la base de datos</i>	43
Figura Nro 13 <i>Preparación de datos para cargar en la base de datos (continuación)</i>	44
Figura Nro 14 <i>Datos epidemiológicos 2019</i>	44
Figura Nro 15 <i>Datos epidemiológicos 2019 (continuación)</i>	45
Figura Nro 16 <i>Datos epidemiológicos 2020</i>	45
Figura Nro 17 <i>Datos epidemiológicos 2020 (continuación)</i>	46
Figura Nro 18 <i>Datos epidemiológicos 2021</i>	46
Figura Nro 19 <i>Datos epidemiológicos 2021 (continuación)</i>	47
Figura Nro 20 <i>Datos epidemiológicos 2022</i>	47
Figura Nro 21 <i>Datos epidemiológicos 2022 (continuación)</i>	48
Figura Nro 22 <i>Datos epidemiológicos 2023</i>	48

Figura Nro 23 <i>Data base del proyecto meteorológico 2019</i>	49
Figura Nro 24 <i>Data base del proyecto meteorológico 2019 (continuación)</i>	49
Figura Nro 25 <i>Data base del proyecto meteorológico 2020</i>	50
Figura Nro 26 <i>Data base del proyecto meteorológico 2020 (continuación)</i>	50
Figura Nro 27 <i>Data base del proyecto meteorológico 2021</i>	51
Figura Nro 28 <i>Data base del proyecto meteorológico 2021 (continuación)</i>	51
Figura Nro 29 <i>Data base del proyecto meteorológico 2022</i>	52
Figura Nro 30 <i>Data base del proyecto meteorológico 2022 (continuación)</i>	52
Figura Nro 31 <i>Data base del proyecto meteorológico 2023</i>	53
Figura Nro 32 <i>Data base del proyecto meteorológico 2023 (continuación)</i>	53
Figura Nro 33 <i>Data base epidemiológico y meteorológico</i>	54
Figura Nro 34 <i>Importaciones</i>	55
Figura Nro 35 <i>Error cuadrático medio (MSE)</i>	58
Figura Nro 36 <i>Error Absoluto Medio (MAE)</i>	58
Figura Nro 37 <i>Interfaz gráfica utilizando tkinter</i>	58
Figura Nro 38 <i>Botón Iniciar Simulación</i>	59
Figura Nro 39 <i>Gráfico de datos reales y predicción</i>	59
Figura Nro 40 <i>Gráfico haciendo predicción 2023 2024</i>	60
Figura Nro 41 <i>Símbolo de sistema sin error</i>	60
Figura Nro 42 <i>Corrección de fecha</i>	62
Figura Nro 43 <i>Prueba de leída de datos reales</i>	62
Figura Nro 44 <i>Prueba de leída de datos reales 2019 a 2023 para predicción del 2024 con problema de fechas justas</i>	63
Figura Nro 45 <i>Prueba de ajuste</i>	63
Figura Nro 46 <i>Visualización de entrenamiento de predicción para el año 2024</i>	64

Figura Nro 47 <i>Prueba de predicción para año 2024</i>	64
Figura Nro 48 <i>Visualización profesional de la simulación</i>	65
Figura Nro 49 <i>Botón selección de año</i>	65
Figura Nro 50 <i>Selección de años futuros</i>	66
Figura Nro 51 <i>Mensaje de advertencia</i>	67
Figura Nro 52 <i>Reinicio de sistema</i>	68
Figura Nro 53 <i>Confirmación para realizar predicción</i>	68
Figura Nro 54 <i>Botón “Ver Casos de Predicción Futura”</i>	69
Figura Nro 55 <i>Ventana de elección de año para ver resultado</i>	69
Figura Nro 56 <i>Ventana de resultados de la simulación</i>	69
Figura Nro 57 <i>Mensaje de error para seleccionar un año</i>	70
Figura Nro 58 <i>Mensaje “No hay predicciones disponibles para el año seleccionado”</i>	70
Figura Nro 59 <i>Botón “Cargar Datos Reales”</i>	71
Figura Nro 60 <i>Selección de datos reales 2024 en archivo Excel</i>	71
Figura Nro 61 <i>Archivo Excel cargado con éxito</i>	72
Figura Nro 62 <i>Botón hacer comparación de casos reales y la predicción</i>	72
Figura Nro 63 <i>Mensaje de error solicitando la carga de los datos reales antes de la comparación</i>	74
Figura Nro 64 <i>Interfaz principal de la simulación</i>	75
Figura Nro 65 <i>Visualización de datos reales del año 2019</i>	75
Figura Nro 66 <i>Visualización de datos reales del año 2020</i>	75
Figura Nro 67 <i>Visualización de datos reales del año 2021</i>	76
Figura Nro 68 <i>Visualización de datos reales del año 2022</i>	76
Figura Nro 69 <i>Visualización de datos reales del año 2023</i>	77
Figura Nro 70 <i>Reinicio del sistema</i>	77

Figura Nro 71 Mensaje para hacer predicción del año seleccionado	78
Figura Nro 72 Predicción del año 2024 mensualmente sin sobre ajuste estudiando los datos históricos epidemiológicos y meteorológicos	79
Figura Nro 73 Visualización de la predicción y la cantidad de contagios en cada mes.....	79
Figura Nro 74 Finalización de la predicción del año 2024	80
Figura Nro 75 Resultado de predicción del año seleccionado.....	80
Figura Nro 76 Datos reales obtenidos del Hospital de Santa Rosa Misiones	81
Figura Nro 77 Visualización de los gráficos casos reales y la predicción con la aparición del cuadro comparativo.....	83
Figura Nro 78 Visualización de los errores Error cuadrático medio (MSE) y Error Absoluto Medio (MAE) con un mensaje calculando que tan precisa es la predicción.....	83
Figura Nro 79 Doble barra invertida para la lectura de la base de datos	87
Figura Nro 80 Función de corrección de fechas de forma automática.....	88
Figura Nro 81 Primera prueba del sistema.....	88
Figura Nro 82 Distribución temporal incorrecta de los casos reales	89
Figura Nro 83 Segunda prueba del sistema	89
Figura Nro 84 Tercera prueba del sistema	90
Figura Nro 85 Cuarta prueba del sistema.....	90
Figura Nro 86 Quinta prueba del sistema.....	91
Figura Nro 87 Modificación del código.....	91
Figura Nro 88 Sexta prueba del sistema	92
Figura Nro 89 Séptima prueba del sistema	92
Figura Nro 90 Simulación de Casos de Dengue: Comparación entre Datos Reales y Predicciones.....	93
Figura Nro 91 Modificación del código.....	93

Figura Nro 92 <i>Casos reales de dengue a lo largo del tiempo en 2022</i>	94
Figura Nro 93 <i>Lectura de datos históricos ordenados desde el año 2019 a 2024</i>	94
Figura Nro 94 <i>Problema de agrupación de fechas con la predicción</i>	95
Figura Nro 95 <i>Visualización de datos reales del año 2019</i>	96
Figura Nro 96 <i>Visualización de datos reales del año 2020</i>	96
Figura Nro 97 <i>Visualización de datos reales del año 2021</i>	97
Figura Nro 98 <i>Visualización de datos reales del año 2022</i>	97
Figura Nro 99 <i>Visualización de datos reales del año 2023</i>	98
Figura Nro 100 <i>Visualización de entrenamiento de predicción para el año 2024</i>	98
Figura Nro 101 <i>Mal resultado</i>	99
Figura Nro 104 <i>Resultado métrica cuadrático absoluto muy alto</i>	99
Figura Nro 103 <i>Modificación del código</i>	99
Figura Nro 104 <i>Modificación del código (continuación)</i>	100
Figura Nro 105 <i>Modificación del código (continuación)</i>	100
Figura Nro 106 <i>Resultado de la modificación del código</i>	100
Figura Nro 107 <i>Visualización del resultado</i>	101
Figura Nro 108 <i>Proceso para integración del archivo Excel</i>	101
Figura Nro 109 <i>Error al cargar los datos reales en Excel</i>	102
Figura Nro 110 <i>Datos reales en archivo Excel modificado</i>	99
Figura Nro 111 <i>Nuevo intento para integración del archivo Excel</i>	102
Figura Nro 112 <i>Lectura del archivo Excel completada sin inconvenientes</i>	103
Figura Nro 113 <i>Precisión del modelo</i>	103

Índice de Tablas

Tabla Nro 1 Comparación de casos reales con la predicción	82
Tabla Nro 2 Comparación de casos reales con la predicción - modificado	104

Resumen

El trabajo final de grado se planteó como objetivo simular dinámicamente en tiempo real el comportamiento de la propagación del dengue a través de Machine Learning, para lo cual buscó identificar posibles deficiencias en el rendimiento del simulador dinámico y evaluar la precisión del mismo. Se aplicó un enfoque mixto, diseño exploratorio e investigación de tipo aplicada, contando con la colaboración de las autoridades de salud del Hospital de Santa Rosa, Misiones, quienes proporcionaron datos meteorológicos suministrados por la Facultad de Ciencias Agrarias correspondientes al período 2019-2024. Se usó el sistema Windows, aplicando el modelo LSTM. Se identificaron y corrigieron varias deficiencias en el rendimiento del simulador, observándose problemas relacionados con el manejo de fechas en el archivo CSV, así como errores de sobreajuste en el modelo. Por otro lado, la precisión del simulador fue evaluada y validada al comparar las predicciones con datos reales de casos de dengue proporcionados por el Hospital de Santa Rosa, Misiones. Los resultados demostraron que el sistema alcanzó una precisión del 95%, con métricas de error MSE y MAE muy bajas, lo que confirma la efectividad del modelo en la predicción de casos de dengue. Para optimizar el simulador y ampliar su aplicabilidad, se propone ampliar el conjunto de datos, explorar modelos complementarios, evaluando la integración de otros modelos de ML junto con LSTM y extender el alcance geográfico.

Palabras clave: simulación dinámica, tiempo real, Machine Learning.

Abstract

The objective of the final degree project was to dynamically simulate in real time the behavior of the spread of dengue through Machine Learning, for which it sought to identify possible deficiencies in the performance of the dynamic simulator and evaluate its accuracy. A mixed approach, exploratory design and applied research was applied, with the collaboration of the health authorities of the Santa Rosa Hospital, Misiones, who will provide meteorological data supplied by the Faculty of Agrarian Sciences corresponding to the period 2019-2024. The Windows system was used, applying the LSTM model. Several deficiencies in the simulator's performance were identified and corrected, observing problems related to the handling of dates in the CSV file, as well as overfitting errors in the model. On the other hand, the accuracy of the simulator was evaluated and validated by comparing the predictions with real data of dengue cases provided by the Santa Rosa Hospital, Misiones. The results showed that the system achieved an accuracy of 95%, with very low MSE and MAE error metrics, confirming the effectiveness of the model in predicting dengue cases. To optimize the simulator and expand its applicability, it is proposed to expand the data set, explore complementary models, evaluating the integration of other ML models together with LSTM and expand the geographical scope.

Keywords: dynamic simulation, real time, Machine Learning.

1. Introducción

El dengue se considera una de las principales enfermedades arbovíricas a nivel mundial, con millones de casos cada año, especialmente en los países tropicales y subtropicales (Ong et al., 2023). La gravedad del problema y la gran presión que ejerce sobre la salud de la población y la economía subrayan la necesidad de desarrollar herramientas efectivas para monitorear y controlar el dengue (Consejo Nacional de Ciencia y Tecnología Paraguay, 2020; Agencia de Información Paraguaya, 2020).

Como se mencionó en el párrafo anterior, la capacidad de predecir y prevenir los brotes de dengue se ha convertido en una de las áreas más prioritarias para los departamentos de salud pública, ya que el virus es transmitido rápidamente con la capacidad de desencadenar una epidemia a gran escala. Por lo tanto, la simulación dinámica en tiempo real del comportamiento de la propagación de la fiebre del dengue a través del uso de las herramientas de Machine Learning (ML) emerge como una técnica innovadora y práctica para el mejoramiento de las operaciones de vigilancia y control de la enfermedad (Ong et al., 2023).

Con la capacidad de aprendizaje de las máquinas, la simulación dinámica en tiempo real es capaz de modelar y predecir la propagación del dengue en diferentes situaciones y entornos, lo que, a su vez, informa a los investigadores y profesionales de la salud pública sobre la implementación de intervenciones de manera más focalizada y exitosa. Varios estudios previos han informado que la modelación matemática es una de las herramientas importantes para predecir brotes y para estimar intervenciones en salud (Consejo Nacional de Ciencia y Tecnología Paraguay, 2020; Agencia de Información Paraguaya, 2020), y ese es el valor del estudio actual en la era moderna.

A diferencia de los modelos tradicionales que operan bajo el supuesto de invarianza temporal, los cuales no consideran la variabilidad en línea, los modelos de ML tienen la capacidad de adaptarse dinámicamente a los flujos de datos en tiempo real. Lo mencionado permite que estos modelos mejoren la precisión de las predicciones y proporcionen información más oportuna y relevante para la toma de decisiones estratégicas (Akhtar et al., 2019).

Sin embargo, uno de los problemas críticos en el desarrollo de simuladores dinámicos en tiempo real es la detección de fallas en el comportamiento del modelo, por lo que la validación de simuladores dinámicos en tiempo real se convierte en un paso crucial para la aplicabilidad efectiva de los simuladores en tiempo real, donde las decisiones tomadas de acuerdo con estos modelos pueden ser cruciales en la influencia de la salud pública. Por lo tanto, en este estudio es de gran utilidad evaluar diferentes algoritmos de ML para determinar el mejor enfoque a utilizar en la predicción del dengue (Roster y Rodrigues, 2021).

La aplicación de ML a la estimación de la propagación del dengue aumenta la precisión de la predicción y se usa para detectar patrones complejos en los datos epidemiológicos. La adaptabilidad de estos algoritmos a nueva evidencia y su capacidad para aprender de volúmenes sustanciales de datos también podrían proporcionar a los investigadores y a los responsables de la toma de decisiones en salud una potente herramienta para la predicción y prevención de brotes (Carvalho, 2020). Como otra contribución, al enfatizar los defectos de los simuladores que ya han sido publicados, este trabajo también intentó mejorar los modelos existentes para permitir el desarrollo de soluciones más sólidas y efectivas para la lucha contra el dengue.

La relevancia de este estudio fue la redefinición de la forma en que se aborda la diseminación del dengue, ya que con la simulación dinámica del comportamiento de la enfermedad se pueden crear escenarios que permitan a los tomadores de decisiones de la esfera de la salud pública tomar decisiones de manera racional y basada en evidencia. Esto es particularmente relevante en un escenario en el que el cambio climático y la urbanización también están alterando los patrones de transmisión del dengue, lo que requiere, más que nunca, del desarrollo de modelos de predicción. El desarrollo de una simulación dinámica que haga uso del Aprendizaje Automático es un activo para la ciencia de la epidemiología y para los profesionales del mundo real que estén manejando datos relacionados con la toma de decisiones en salud pública.

En este escenario, en este trabajo final de grado se definió un diseño de simulador dinámico en tiempo real para modelar la difusión del dengue en el departamento de Misiones utilizando técnicas de ML muy avanzadas. La falta de cualquier tipo de modelo de simulación en tiempo real en este campo de estudio ha llevado a la escasez de información muy importante que se puede utilizar para implementar estrategias de

prevención y control efectivas. Por lo tanto, este proyecto de investigación tuvo como objetivo solucionar ese problema creando un simulador que tenga la capacidad de proporcionar información en tiempo real sobre los casos probables de ser positivos para la enfermedad del dengue. El simulador se escribió usando Python como lenguaje de programación porque tiene un gran conjunto de bibliotecas relacionadas con ML, como scikit-learn, TensorFlow y Keras, y también porque su capacidad de prototipado es muy rápida.

El desarrollo del simulador siguió una serie de etapas clave. La primera fue la adquisición de datos epidemiológicos, ambientales y sociales de fuentes confiables como hospitales, clínicas, oficinas gubernamentales y estaciones meteorológicas. Los datos estuvieron sujetos a un proceso de preprocesamiento para limpiarlos y afinarlos para su análisis, lo cual incluyó la corrección de errores, la eliminación de valores atípicos, la imputación de datos faltantes, y luego la normalización y transformación de los datos para hacerlos uniformes y en línea con los algoritmos de ML.

Una vez que los datos fueron sometidos a algún nivel de procesamiento previo, el siguiente paso fue la construcción del modelo de simulación, en el que se eligieron los algoritmos pertinentes para representar la naturaleza de la propagación del dengue, y se entrenó el modelo con los datos de entrenamiento disponibles. El modelo se integró en un entorno de simulación en tiempo real, a través del cual se alimentó constantemente con nuevos datos e hizo nuevas predicciones sobre la propagación del dengue en tiempo real de manera dinámica.

El rendimiento del simulador se siguió y actualizó regularmente para garantizar la máxima precisión y fiabilidad, y se probó a fondo su robustez y capacidad de respuesta en todas las situaciones y entornos. Además, tanto el desarrollo del simulador como sus detalles de rendimiento estuvieron basados en casos de uso específicos, para los cuales se recopilaron datos del mundo real y se analizaron los brotes de dengue en el Departamento de Misiones. Luego, el rendimiento del simulador se validó y verificó a través de la comprobación de las predicciones realizadas por el modelo con datos de la vida real, y se utilizaron métricas de evaluación para medir la confiabilidad y robustez de los resultados.

Una vez validado, el simulador se implementó para su uso en el Departamento de Misiones y, posiblemente, en otros departamentos y regiones afectadas por el dengue. Para este propósito, se diseñó una interfaz de usuario fácil de usar que permitió a los usuarios interactuar con el simulador y extraer información relevante de manera sencilla y rápida.

Atendiendo a todo lo mencionado anteriormente, se comprobaron los resultados derivados del simulador para obtener conclusiones sobre la eficacia del programa de prevención y control del dengue. Los resultados y las recomendaciones sobre lo que debe o no hacerse se presentaron al construir e implementar el simulador.

1.1. Planteamiento del problema

El dengue constituye una enfermedad arboviral con potencial letal, la cual es transmitida por mosquitos hembra del género *Aedes* (Cabrera, 2019; Castrillón et al., 2015), siendo *A. aegypti*, *A. albopictus* y *A. vittatus* las especies más relevantes, estos vectores se caracterizan por ser ectoparásitos hematófagos comunes en áreas tropicales. Esta enfermedad zoonótica se propagó desde primates no humanos originarios de África o Asia hace entre 500 y 1000 años (Holmes y Twiddy, 2003). No obstante, en las últimas seis décadas, su diseminación ha evolucionado de un escenario en el que solo nueve países enfrentaban epidemias severas a convertirse en endémica en más de 100 países a nivel mundial, afectando incluso a regiones no tropicales o subtropicales (Messina et al., 2014; Saydam et al., 2021). Adicionalmente, se estima que aproximadamente cien millones de individuos padecen anualmente la enfermedad sintomática (Messina et al., 2014), la cual es provocada por sus cuatro serotipos (Iqbal e Islam, 2017).

Ante el considerable impacto que los cambios ambientales ejercen sobre la transmisión de esta enfermedad, se hace imprescindible la implementación urgente del enfoque de una salud que aboga por la integración de la salud humana, animal y ecológica (Cabrera et al., 2022). Los factores tradicionales vinculados al dengue abarcan condiciones sociales, económicas, ambientales, culturales, socioeconómicas, la pobreza y aspectos relacionados con el bienestar de la población (Cabrera y Taylor, 2019). Sin embargo, el desarrollo de las larvas de mosquitos y las condiciones climáticas son algunos de los elementos influyentes que requieren una investigación más exhaustiva (Jacome et al., 2019).

Los especialistas en epidemiología sostienen que la comprensión de los ciclos de vida de las especies relevantes en sus hábitats naturales, en contraste con las condiciones experimentales, representa un desafío científico que debe ser superado para avanzar hacia la creación de una iniciativa más eco-inspirada para el control y la prevención de enfermedades infecciosas emergentes y reemergentes. En consecuencia, se requiere una comprensión más profunda del entorno local (Destoumieux-Garzón et al., 2018).

Aunque existe un amplio debate sobre la relevancia del cambio climático en la distribución espacial de las enfermedades infecciosas a nivel global (Destoumieux-Garzón et al., 2018; Zambrano et al., 2017), los efectos rezagados también son una característica dependiente del contexto local (Cabrera y Taylor, 2019) que necesita ser explicada con mayor claridad.

A nivel internacional, el dengue emergió como una de las mayores amenazas a la salud pública, con aproximadamente 390 millones de infecciones por año en más de 100 países. El hiperdesarrollo de las ciudades y el cambio climático, además de vacíos en medidas eficientes de control de vectores, han intensificado la expansión del virus transmitido por mosquitos, conduciendo a un alarmante aumento tanto en términos de frecuencia como de gravedad de la enfermedad en varias partes del mundo. Este panorama internacional enfatiza el ciclo de necesidad previamente mencionado, donde nuevas herramientas efectivas e innovadoras deben desarrollarse y utilizarse para prever y, finalmente, prevenir la expansión del dengue, especialmente en áreas vulnerables con pocas posibilidades de proporcionar apoyo en términos de recursos sanitarios.

Hasta la fecha, América Latina y el Caribe ha reportado un total de 9,3 millones de casos de dengue en 2024; esta cifra es más del doble de los casos registrados en 2023 y subraya la importancia de mantener la vigilancia, intensificar las medidas de prevención y control de emergencia, y garantizar el acceso oportuno a la atención médica.

Hasta la semana epidemiológica 16 de 2024, los países de América Latina y el Caribe han reportado más de 9500 casos de dengue grave, lo que se tradujo en más de 4500 muertes con una tasa de letalidad del 0,048%. Comparativamente, con el mismo periodo de 2023, se ha observado un aumento en el número de casos en todos los países. Este aumento puede atribuirse a varios factores, incluida la expansión territorial de su principal vector, el mosquito *Aedes aegypti*, que ha extendido su área de distribución a regiones donde

previamente no existía presencia de este vector, posiblemente como resultado de fenómenos como El Niño y el cambio climático (PAHO, 2024).

Por otro lado, la expansión urbana rápida, no planificada, y el aumento de la población urbana, así como la asequibilidad limitada de los servicios de agua y saneamiento, han propiciado recipientes en donde los mosquitos depositan sus huevos, lo cual también contribuye a la proliferación de mosquitos. Además, la menor exposición humana al dengue, debido a la reducción de la movilidad y la interacción social entre humanos y mosquitos, principalmente causada por el confinamiento y las restricciones a los viajes y movilidad forzada por la epidemia de COVID-19, puede haber aumentado la población susceptible (PAHO, 2024).

La PAHO (2024), por otro lado, ha informado que con la temporada de circulación de dengue en el norte del hemisferio aumentando rápidamente, la OPS está llamando a todos los países y territorios a intensificar sus esfuerzos para la detección y prevención de casos. Asimismo, señala que los gobiernos también deben concienciar a la comunidad sobre la necesidad de eliminar los criaderos de mosquitos, aplicar medidas de protección personal para prevenir las picaduras de mosquitos y promover una atención médica rápida y adecuada para los pacientes.

Sin embargo, Paraguay ha experimentado brotes recurrentes –y, en la mayoría de los casos, incrementados– de dengue, que han tenido un severo impacto en su población. Aunque las autoridades de salud han hecho lo que está a su alcance para predecir y prevenir rebrotes de la enfermedad, esto ha demostrado ser insuficiente. En realidad, mientras que, en un país como Paraguay, previamente se ha demostrado que los modelos predictivos basados en los datos son prometedores, la recolección y análisis de datos a nivel local es un área que aún enfrenta desafíos (McMurren et al., 2017). Por lo tanto, considerando que el dengue es común y se ha identificado un aumento significativo de su incidencia en el departamento de Misiones, la incorporación de datos epidemiológicos, ambientales y sociales en un modelo dinámico en tiempo real de infección podría ser una pieza vital para mejorar el pronóstico y manejo de las enfermedades en tal región.

San Juan, San Ignacio y Santa Rosa, tres distritos en el Departamento de Misiones, tenían los mayores picos de pacientes con dengue, respectivamente, siendo declarados zona roja (Rodríguez, 2024). Por consiguiente, extraídos los últimos datos, aumentó de manera significativa el número de casos positivos de la enfermedad en Misiones, notificándose muertes por la misma enfermedad. En este contexto, Riveros (2024) menciona que la situación epidemiológica en el departamento se encuentra complicada, y parece haber mucho vacío de acciones o implementadas incorrectamente. Se observa hoy la falta de modelos de simulación en tiempo real capaces de informar de forma precisa y oportuna sobre la propagación de la enfermedad, y esto ha llevado a una respuesta limitada de las autoridades sanitarias. La falta de integración de datos y análisis predictivo de esta región subrayó la importancia de desarrollar urgentemente un simulador dinámico en tiempo real controlado por técnicas de Aprendizaje Automático avanzadas para modelar la propagación del dengue, puesto que, además de hacer que sea posible predecir la incidencia futura de esta enfermedad, también sería de gran valor como herramienta para mejorar la toma de decisiones en materia de salud pública y la implementación de estrategias de control más eficientes en la región de Misiones.

1.2. Justificación

El dengue se define como una enfermedad viral transmitida por mosquitos, que representa un grave desafío para la salud pública a nivel global. De acuerdo con la Organización Mundial de la Salud (OMS), se estima que anualmente se producen entre 100 y 400 millones de infecciones por dengue, de las cuales solo una pequeña proporción se manifiesta clínicamente, siendo la rápida expansión del dengue debida a diversos factores como el cambio climático, los viajes internacionales, la urbanización descontrolada y la falta de un control apropiado sobre el vector transmisor.

En este contexto, resulta fundamental simular la dinámica de propagación en tiempo real del dengue para poder anticipar y disminuir el impacto de esta enfermedad, siendo los modelos de simulación desarrollados por medio del Aprendizaje Automático los que permiten introducir las múltiples variables que determinan la acción del virus, tales como las condiciones del clima, los patrones de movilidad humana, las características sociodemográficas, entre otros. Así, la simulación de las dinámicas mencionadas,

realizadas por estos modelos, permite detectar los focos de infección, anticiparse a los brotes epidémicos y valorar la efectividad de las estrategias de minimización.

Ante lo expuesto, la presente investigación se justificó por la necesidad de desarrollar modernas herramientas de simulación para mejorar la comprensión de la propagación del dengue y apoyar la toma de decisiones en salud pública, puesto que, generando un simulador dinámico en tiempo real, basado en ML, se pueden encontrar lagunas en el rendimiento y evaluar la precisión de las predicciones. La posibilidad de modular rápidamente e iterativamente el modelo conduce a un ciclo de mejora continua, adaptado a las circunstancias específicas de cada geografía y epidemiología.

Presentados de esta manera, los resultados de la investigación pueden resultar de gran utilidad para las autoridades sanitarias y los responsables de la toma de decisiones, ya que funcionan como un sistema de alerta temprana y un laboratorio virtual en el que es posible evaluar estrategias de control del dengue. Además, la metodología empleada y las lecciones aprendidas durante el desarrollo de la investigación son susceptibles de ser replicadas en el estudio de otras enfermedades transmitidas por vectores, como el zika o el chikungunya.

1.3. Alcance y Limitaciones

Dentro del ámbito temático, se desarrolló un simulador dinámico en tiempo real del comportamiento de la propagación del dengue utilizando técnicas de Aprendizaje Automático. El enfoque estuvo en identificar y predecir la propagación del dengue en el Departamento de Misiones, así como en proporcionar información inmediata sobre posibles casos positivos.

En cuanto al marco temporal, el proyecto se llevó a cabo a lo largo del año 2024.

Geográficamente, la investigación se centró en el Departamento de Misiones, con un enfoque específico en las áreas afectadas por la propagación del dengue.

Desde un enfoque metodológico, se emplearon métodos de modelado y modelado de máquinas para desarrollar el simulador. Se utilizaron técnicas de adquisición y preprocesamiento de datos para recopilar y preparar datos relevantes sobre el dengue; además, se implementaron modelos de Aprendizaje Automático para predecir la propagación de la enfermedad.

En cuanto a los participantes, se contó con la colaboración de las autoridades de salud del Hospital de Santa Rosa, Misiones, quienes proporcionaron datos meteorológicos suministrados por la Facultad de Ciencias Agrarias correspondientes al período 2022-2024, lo cual permitió disponer de información actualizada para compararla con los datos obtenidos en el estudio.

Tecnológicamente, el simulador tiene la capacidad de automatizar la recolección y el análisis de datos, lo que disminuye la carga de trabajo manual y optimiza la eficiencia en el proceso de toma de decisiones. Las autoridades sanitarias podrán recibir alertas automáticas y recomendaciones fundamentadas en datos actualizados, lo que promueve una gestión más proactiva de los recursos y de las medidas de control.

El sistema de simulación estuvo constituido por seis módulos: la adquisición de datos, que se encargó de reunir información fundamental sobre el dengue, incluyendo casos confirmados y datos climáticos, garantizando la calidad y precisión necesarias; el preprocesamiento de datos, que se dedicó a limpiar, transformar y normalizar la información para prepararla para el análisis, eliminando errores y estandarizando formatos; el Aprendizaje Automático, donde se implementaron algoritmos de ML para anticipar la propagación del dengue, ajustando los modelos para captar la complejidad de la enfermedad; la simulación dinámica, que integró estos modelos en un entorno en tiempo real, permitiendo que el sistema se adapte y actualice continuamente con nuevos datos; la evaluación y validación, que comparó las predicciones con datos reales para medir la precisión y robustez del simulador; y, por último, el despliegue y diseminación, que ofreció una interfaz de usuario intuitiva, facilitando el acceso a la información y su aplicación en la toma de decisiones en el ámbito de la salud pública.

Cabe destacar que se usó el sistema Windows debido a la amplia compatibilidad que este ofrece con diversas herramientas de desarrollo y bibliotecas necesarias para implementar redes neuronales y la visualización de datos epidemiológicos, proporcionando un entorno estable y familiar que facilita la integración de software como Python y Pandas.

En lo que refiere a las limitaciones, se encontró la complejidad intrínseca de los modelos de Aprendizaje Automático aplicados a la salud pública, puesto que, a pesar de que se emplearon técnicas avanzadas para modelar la propagación del dengue, el comportamiento de la enfermedad puede verse afectado por una diversidad de factores que no se reflejan de manera integral en el modelo. En este contexto, la capacidad del simulador para ajustarse a variaciones en las condiciones ambientales o sociales es restringida, lo que podría mermar su efectividad en situaciones imprevistas.

Por otro lado, la recolección y utilización de datos epidemiológicos, ambientales y sociales pueden incluir información personal sensible, como datos de salud o información demográfica pormenorizada, por lo cual fue fundamental garantizar que tales datos sean gestionados bajo rigurosas normas de confidencialidad, dado que la carencia de un manejo apropiado de esta información pone en riesgo la privacidad de los individuos y puede dar lugar a conflictos éticos.

Desde el punto de vista tecnológico, para que el simulador opere en tiempo real, fue fundamental que los datos se procesaran y analizaran de forma inmediata. No obstante, el sistema presentó una latencia mínima de aproximadamente 10 segundos por predicción, lo cual se considera dentro de los parámetros normales.

1.4. Objetivos

1.4.1. Objetivo general

Simular dinámicamente en tiempo real el comportamiento de la propagación del dengue a través de Machine Learning.

1.4.2. Objetivos específicos

Identificar posibles deficiencias en el rendimiento del simulador dinámico del comportamiento de la propagación del dengue.

Evaluar la precisión del simulador dinámico desarrollado en la predicción de la propagación del dengue.

1.5. Contenido del libro

El presente trabajo de investigación simuló dinámicamente en tiempo real el comportamiento de la propagación del dengue a través de Machine Learning, estructurándose de la siguiente forma:

- **Introducción:** Se introduce el tema de la investigación, proporcionando una visión general del problema del dengue a nivel mundial, regional y local, con un enfoque en la importancia de desarrollar un simulador dinámico en tiempo real utilizando Machine Learning. Se presentan los objetivos generales y específicos del estudio, y se detalla la estructura del trabajo.
- **Marco Teórico:** Se realiza un análisis exhaustivo del estado del arte en simulación de enfermedades infecciosas, especialmente del dengue, utilizando Machine Learning. Se revisan conceptos fundamentales, estudios previos y las bases teóricas que sustentan el desarrollo del simulador propuesto.
- **Marco Metodológico:** Se describe la metodología utilizada para desarrollar el simulador, incluyendo la recolección y preprocesamiento de datos, la selección de algoritmos de Machine Learning y la implementación del simulador en un entorno de simulación en tiempo real. Se detallan las herramientas y técnicas empleadas, así como los criterios de evaluación del modelo.

- **Conclusiones:** Se mencionan los principales hallazgos y conclusiones de la investigación, evaluando si se alcanzaron los objetivos propuestos. Se discuten las implicaciones prácticas del simulador para la prevención y control del dengue, y se ofrecen recomendaciones para su implementación y mejora.
- **Trabajos Futuros:** Se identifican posibles áreas de investigación futuras, sugiriendo extensiones del modelo y nuevas aplicaciones del simulador en otros contextos o enfermedades.
- **Anexos:** Se incluyen anexos con información adicional, como detalles técnicos, tablas de datos y gráficas de simulación, que complementan la discusión principal del trabajo.
- **Glosario de Términos y Abreviaturas:** Se presenta un glosario que define los términos técnicos y abreviaturas utilizadas en el documento.
- **Referencias Bibliográficas:** Se listan todas las fuentes y referencias utilizadas, siguiendo las normas APA séptima edición, para respaldar las afirmaciones y argumentos del trabajo.

aprobada únicamente para niños y adolescentes que tengan un historial de infección confirmada por el virus del dengue a través de pruebas de laboratorio (Yuill, 2023).

Una investigación llevada a cabo por Gubler (2019), muestra que las tasas de incidencia del dengue han aumentado significativamente en las últimas décadas, con brotes cada vez más frecuentes en áreas que antes no habían sido afectadas. Este fenómeno ha sido impulsado por la globalización, el cambio climático y la urbanización descontrolada (Gubler, 2019). Un análisis de Shepard et al. (2020) revela que el dengue afecta a más de 100 países y que cerca de la mitad de la población mundial está en riesgo de contraer la enfermedad. En particular, los países de América Latina y el sudeste asiático son los más afectados. Los ciclos epidémicos del dengue tienden a repetirse cada 3 a 5 años en estas áreas, de acuerdo con la inmunidad de la población y la circulación de los 4 serotipos del virus (Shepard et al., 2020).

El impacto económico del dengue también es considerable, no solo por los gastos directos relacionados con el tratamiento médico, sino también por la pérdida de productividad laboral. Un estudio reciente de Bhatt et al. (2021) estimó que los costos globales asociados al dengue superan los 8.9 mil millones de dólares anuales, lo que resalta la necesidad de implementar intervenciones más efectivas para su control. Este estudio también subrayó la importancia de invertir en vacunas y nuevas tecnologías para el control de vectores, como los mosquitos genéticamente modificados que buscan reducir la transmisión (Bhatt et al., 2021). El control del dengue sigue siendo un desafío a nivel global. Estrategias como la vigilancia epidemiológica, el control de vectores y la educación comunitaria son fundamentales para la prevención de brotes. No obstante, Halstead (2020) señala que la implementación de estas medidas varía considerablemente entre los diferentes países, lo que reduce su efectividad. Halstead también subraya la importancia de programas integrados que adopten los esfuerzos de salud pública junto a la investigación científica para el desarrollo de vacunas más accesibles y efectivas (Halstead, 2020).

2.2.2. Impacto del dengue en la salud pública del Paraguay

El dengue ha generado un impacto significativo en la salud pública de Paraguay en las últimas décadas, convirtiéndose en una de las enfermedades que requieren notificación obligatoria en el país. De acuerdo con información del Ministerio de Salud Pública y Bienestar Social de Paraguay, el país ha enfrentado varios brotes importantes, siendo los

más severos los ocurridos en 2013, 2016 y 2020, con una alta tasa de incidencia que ha puesto a prueba el sistema de salud pública (Ministerio de Salud Pública y Bienestar Social, 2020). Un análisis llevado a cabo por Arbo y Paredes (2019) subraya que los elementos que favorecen la difusión del dengue en Paraguay abarcan el rápido crecimiento de las áreas urbanas, la insuficiencia de infraestructura sanitaria adecuada y las condiciones climáticas propicias para la reproducción del mosquito *Aedes aegypti*. Esta combinación de factores, junto con la ineficacia en el control del vector, ha incrementado la vulnerabilidad de la población frente a la enfermedad (Arbo & Paredes, 2019). El efecto económico del dengue en Paraguay ha sido considerable. De acuerdo con un estudio de Ortiz et al. (2021), los gastos relacionados con el tratamiento de pacientes con dengue en el país han crecido de manera notable, no solo debido a los recursos médicos requeridos durante los brotes, sino también por la pérdida de productividad laboral (Ortiz et al., 2021). De manera similar, la presión sobre los servicios de salud pública durante los brotes de dengue ha puesto de manifiesto la debilidad del sistema sanitario en Paraguay; en momentos de picos epidémicos, los hospitales y centros de atención primaria se encuentran desbordados, lo que impacta negativamente la capacidad del sistema para atender otras enfermedades y emergencias. Como indica un estudio de Giménez et al. (2020), la carencia de infraestructura adecuada y la falta de personal sanitario capacitado han sido factores determinantes en la ineficaz respuesta a los brotes.

A pesar de los intentos del gobierno paraguayo y las campañas de sensibilización, el dengue continúa representando un riesgo para la salud pública en el país. La ausencia de una vacuna accesible y el aumento de la resistencia de los mosquitos a los insecticidas convencionales dificultan las estrategias de control. De acuerdo con la Organización Panamericana de la Salud (2021), es fundamental adoptar un enfoque integral que una el control del vector, la educación de la comunidad y la vigilancia epidemiológica para reducir el impacto del dengue en Paraguay (OPS, 2021). En Paraguay, la tasa de mortalidad asociada al dengue ha generado una creciente inquietud, especialmente durante los brotes de los últimos años. De acuerdo con un informe del Ministerio de Salud Pública y Bienestar Social (2021), el brote de 2020 fue el más grave registrado en la historia del país, con más de 200,000 casos notificados y una tasa de letalidad del 0.12%, lo que representó un aumento en comparación con años anteriores. Este aumento se atribuyó a una mayor circulación del serotipo DEN-2, que ha estado relacionado con casos

más graves y un mayor riesgo de complicaciones. Las autoridades de salud han enfatizado la importancia de fortalecer los sistemas de atención primaria para identificar y tratar los casos de dengue de manera oportuna, con el objetivo de prevenir su avance hacia formas severas de la enfermedad (Ministerio de Salud Pública y Bienestar Social, 2021).

Otro elemento importante es la susceptibilidad de las áreas urbanas con alta densidad poblacional, como Asunción y Ciudad del Este, donde los elevados índices de pobreza y las malas condiciones de saneamiento han propiciado la proliferación del mosquito vector. Estudios recientes, como el de González et al. (2022), indican que las políticas públicas deberían concentrarse no solo en la respuesta sanitaria, sino también en mejorar las condiciones socioeconómicas de estas zonas para disminuir los criaderos de mosquitos. Las campañas de sensibilización comunitaria, la participación activa de la población en la eliminación de criaderos y la implementación de medidas sostenibles para el control del vector han demostrado ser estrategias efectivas, aunque su aplicación ha sido desigual en distintas regiones del país (González et al., 2022).

2.2.3. Modelos Matemáticos y Computacionales de propagación de enfermedades

Los modelos matemáticos han sido herramientas esenciales para comprender la dinámica de la propagación de enfermedades infecciosas; entre ellos, el modelo SIR (Susceptible-Infectado-Recuperado) se destaca como uno de los más utilizados, el cual organiza a la población en tres categorías: susceptibles (S), infectados (I) y recuperados (R), y utiliza ecuaciones diferenciales para describir las tasas de cambio entre estos grupos. Otros modelos relacionados son el SEIR (Susceptible-Expuesto-Infectado-Recuperado), que incorpora una fase de exposición previa a la infección activa, y el SIS (Susceptible-Infectado-Susceptible), que considera la posibilidad de reinfecciones tras la recuperación. Los modelos mencionados son fundamentales para anticipar cómo puede propagarse una enfermedad y qué medidas podrían implementarse para controlar la epidemia (Anderson y May, 1991).

Aparte de los modelos señalados, existen variantes que añaden mayor complejidad, como el modelo SEIRS, que incluye una fase de pérdida de inmunidad, y el modelo MSEIR, que considera una etapa de protección inmunológica materna, siendo cruciales al modelar enfermedades con características particulares, como la posibilidad de reinfección o inmunidad temporal, y proporcionando un marco más sólido para realizar pronósticos

(Brauer y Castillo-Chávez, 2012). La capacidad de estos modelos para ser ajustados y ampliados los hace herramientas valiosas para simular epidemias bajo diferentes circunstancias y escenarios. Por otro lado, los modelos metapoblacionales, que dividen a la población en subgrupos conectados mediante movilidad o intercambio de individuos, permiten estudiar la propagación de enfermedades en contextos geográficos más realistas; son especialmente útiles en la simulación de la diseminación de enfermedades en zonas urbanas y rurales, donde la conectividad y la densidad poblacional juegan un papel fundamental (Grenfell y Harwood, 1997).

2.2.3.1. Simulación dinámica

Constituye un método computacional que facilita el modelado del comportamiento de sistemas complejos a lo largo del tiempo, siendo particularmente valiosa en la propagación de enfermedades infecciosas. Por medio de la simulación dinámica, se pueden examinar diversos escenarios y valorar el impacto de diferentes estrategias de intervención. Investigaciones anteriores han utilizado simulaciones dinámicas para comprender cómo las variaciones en la movilidad de la población o la implementación de cuarentenas afectan la propagación de enfermedades como el dengue o la gripe (Keeling & Rohani, 2008). Sin lugar a dudas, este tipo de simulaciones permite realizar pronósticos más precisos al considerar la variabilidad y la estocasticidad en los modelos, algo que los modelos matemáticos deterministas no logran hacer por sí solos.

Por su lado, la simulación en tiempo real es un método que permite llevar a cabo modelos simulados de tal manera que los resultados se presenten conforme se generan, reproduciendo la evolución de un sistema en tiempo real. Resulta crucial para aplicaciones en las que la rapidez de respuesta es vital, como en la administración de brotes epidémicos, donde las decisiones deben fundamentarse en datos actualizados y precisos (Gordon et al., 2015). La utilización de simulaciones en tiempo real permite integrar datos en vivo en el modelo, lo que mejora la exactitud y pertinencia de las predicciones. Esto, en el contexto del dengue, donde factores como el clima y la movilidad de la población pueden cambiar de forma rápida y afectar la dinámica del brote, es sumamente importante. Así, a través de la actualización constante del modelo con datos en tiempo real, se pueden adaptar las estrategias de control de manera más eficaz, optimizando la utilización de recursos y mejorando los resultados en salud (Perkins et al., 2019).

Asimismo, la simulación dinámica es útil para analizar el efecto de intervenciones no farmacéuticas, como el distanciamiento social, las campañas de vacunación y las medidas de control vectorial en el caso del dengue. Mediante la simulación, se pueden explorar diferentes estrategias y combinaciones de medidas, permitiendo que los responsables de políticas de salud tomen decisiones informadas sobre las mejores acciones a seguir para mitigar o controlar un brote (Ferguson et al., 2006), lo cual es especialmente relevante en enfermedades como el dengue, donde la disponibilidad de vacunas es limitada y el control vectorial es esencial para la prevención. Un aspecto fundamental de las simulaciones dinámicas es la capacidad de ajustar y calibrar los modelos en función de datos en tiempo real, lo que mejora significativamente la exactitud y relevancia de las predicciones. Las simulaciones pueden ser continuamente perfeccionadas a medida que se reciben nuevos datos, permitiendo que los modelos se desarrollen y se adapten a la situación actual del brote (Merler et al., 2015).

La ejecución de simulaciones en tiempo real demanda el uso de técnicas y herramientas específicas que permitan manejar grandes cantidades de datos y procesarlos con rapidez, siendo las más frecuentes los sistemas de computación en paralelo y las arquitecturas de alta disponibilidad, que posibilitan que el simulador procese la información en tiempo real sin demoras significativas (Eubank et al., 2004). También, la optimización de los algoritmos es esencial para asegurar que las simulaciones se realicen en un periodo de tiempo reducido sin sacrificar la precisión, abarcando el uso de técnicas como el balanceo de carga y la optimización del código, que garantizan que los recursos computacionales se utilicen de manera eficaz y que los resultados se obtengan con la rapidez necesaria para la toma de decisiones. Otro aspecto significativo en la implementación de simulaciones en tiempo real es la integración de sistemas de monitoreo y alerta temprana, que pueden identificar cambios importantes en la dinámica del brote y activar respuestas automáticas dentro del modelo. Estos permiten no solo la simulación, sino también la gestión activa del brote en tiempo real, facilitando la coordinación de intervenciones en múltiples frentes, desde el control de vectores hasta la comunicación de riesgos a la población (Lorch et al., 2020).

2.2.3.2. Machine Learning en la simulación de enfermedades

El Machine Learning ha emergido como una herramienta eficaz para mejorar la modelación y la previsión de la propagación de enfermedades, puesto que, en lugar de depender únicamente de ecuaciones predeterminadas, los modelos de Machine Learning aprenden a identificar patrones directamente a partir de los datos, lo que permite una mayor adaptabilidad y precisión. Algoritmos como las redes neuronales, los árboles de decisión y las máquinas de soporte vectorial (SVM) han sido aplicados con éxito en la anticipación de brotes de enfermedades y en la optimización de simulaciones epidemiológicas. En este contexto, investigaciones anteriores han mostrado que la integración del ML en simulaciones epidemiológicas puede aumentar de manera significativa la exactitud de las previsiones, permitiendo respuestas más rápidas y efectivas frente a brotes epidémicos (Zou et al., 2020).

La utilización del Machine Learning en la simulación de enfermedades ha permitido gestionar grandes cantidades de datos de manera efectiva, lo que resulta fundamental en la era de los big data. Con el crecimiento acelerado de la información disponible, como los registros electrónicos de salud, datos de movilidad y redes sociales, los modelos de Machine Learning pueden extraer información pertinente y actualizar las predicciones de forma instantánea. En enfermedades con patrones de propagación complicados, esto es particularmente valioso, donde múltiples factores, incluyendo el clima, la movilidad de las personas y las intervenciones en salud, influyen en la dinámica de la epidemia (Dai et al., 2016). También, los modelos de Machine Learning han permitido una mayor adaptación en las simulaciones, ajustando las predicciones no solo a nivel de la población en general, sino también a nivel individual o de subgrupos. Esto ha llevado a un enfoque más detallado en la anticipación de brotes, donde se pueden identificar subgrupos en alto riesgo y dirigir intervenciones específicas. Como, por ejemplo, en el contexto del dengue, el Machine Learning puede ayudar a prever la probabilidad de brotes en barrios concretos o durante ciertos periodos climáticos, mejorando así la efectividad de las medidas de control vectorial (Salje et al., 2016).

2.2.4. Evaluación del rendimiento de simulaciones dinámicas

2.2.4.1. Deficiencias en el rendimiento de un simulador dinámico

Una técnica habitual para detectar deficiencias es el análisis de residuos, que examina las discrepancias entre las predicciones del modelo y los datos reales para identificar patrones sistemáticos de error, patrones que pueden indicar la existencia de sesgos en el modelo o la necesidad de incluir factores adicionales que no se habían considerado previamente (Saltelli et al., 2004). De acuerdo con Melo-González (2006), los simuladores dinámicos se fundamentan en modelos matemáticos que pueden presentar limitaciones al intentar reflejar la complejidad de los sistemas reales. Estos modelos simplificados podrían no abarcar todos los fenómenos significativos, lo que resulta en pronósticos inexactos sobre el rendimiento. En cambio, Luyben (2002) indica que la calidad de los datos de entrada, tales como las propiedades de los materiales y las condiciones de operación, influye de manera directa en la exactitud de los resultados de la simulación. La falta de certeza en estos datos puede extenderse e intensificarse, lo que lleva a pronósticos poco fiables sobre el rendimiento.

Por su parte, Svrcek et al. (2006) mencionan que los algoritmos numéricos empleados en los simuladores dinámicos pueden enfrentar dificultades de convergencia, particularmente en sistemas que presentan comportamientos muy no lineales o inestables. Esto puede resultar en resultados incorrectos o en la imposibilidad de alcanzar una solución. También, Marlin (2005) subraya la relevancia de validar los modelos de simulación utilizando datos experimentales. Sin una validación adecuada, no se puede tener confianza en que las predicciones del simulador reflejen el rendimiento real del sistema. La carencia de datos de validación es una limitación frecuente en numerosos estudios de simulación.

2.2.4.2. La precisión, eficiencia y robustez del sistema

La valoración del desempeño de simulaciones dinámicas en tiempo real se efectúa a través de métricas que toman en cuenta tanto la precisión de las predicciones como la eficiencia computacional del sistema, considerando el tiempo de ejecución, la latencia en el procesamiento y la capacidad para gestionar grandes volúmenes de datos. Aspectos como la capacidad de procesamiento disponible y la calidad de los datos de entrada pueden impactar de manera notable en estos parámetros, por lo que deben ser cuidadosamente

optimizados (Kiss et al., 2017). La precisión de las predicciones se analiza comparando los resultados de la simulación con datos observados, aplicando métricas como el error cuadrático medio (MSE) y la tasa de aciertos.

Además de la precisión y la eficiencia, otro factor esencial en la evaluación del rendimiento es la robustez del sistema, es decir, su aptitud para manejar variaciones en los datos de entrada y en las condiciones del entorno sin que el rendimiento se vea afectado de manera significativa, lo cual resulta particularmente importante en simulaciones en tiempo real, donde las condiciones pueden cambiar velozmente y el sistema debe ser capaz de ajustarse y continuar ofreciendo resultados confiables (Balcan et al., 2009).

2.25. Algoritmos de Machine Learning

Suelen clasificarse en tres grandes grupos: aprendizaje supervisado, no supervisado y por refuerzo. En el caso del aprendizaje supervisado, se emplean datos que han sido etiquetados para entrenar al modelo, lo que le permite hacer predicciones sobre nuevos datos. Dentro de esta categoría, los algoritmos de regresión y clasificación son representativos. Las redes neuronales forman parte de esta clasificación, las cuales son apropiadas para modelar relaciones complejas y no lineales en los datos epidemiológicos, mientras que los árboles de decisión proporcionan una interpretación clara de las decisiones del modelo. Por otro lado, las máquinas de soporte vectorial (SVM) son eficaces para la clasificación y predicción en situaciones donde los datos son limitados o ruidosos (Bishop, 2006).

Específicamente, las redes neuronales profundas han demostrado ser efectivas en la detección de patrones complejos en grandes conjuntos de datos, lo cual es particularmente útil en la simulación del dengue, donde múltiples variables como la temperatura, la humedad y la densidad poblacional interactúan de forma no lineal (LeCun et al., 2015). Por otro lado, el aprendizaje no supervisado se ocupa de datos que no tienen etiquetas, intentando identificar patrones o agrupamientos en la información, como ocurre en el clustering. Por último, el aprendizaje por refuerzo consiste en que un agente adquiere la habilidad de tomar decisiones a través de su interacción con un entorno y la retroalimentación obtenida de sus acciones, mejorando su comportamiento con el tiempo (IBM, 2024). Las aplicaciones del ML son amplias y cubren diversos sectores. En el

campo del comercio, se implementan algoritmos para ofrecer recomendaciones personalizadas de productos, tomando en cuenta el comportamiento del consumidor. En el ámbito de la salud, el aprendizaje automático se utiliza para realizar diagnósticos automáticos a partir de imágenes médicas, lo que contribuye a aumentar la precisión y velocidad en la identificación de enfermedades. Asimismo, en el sector financiero, se aplica para detectar fraudes y para evaluar los riesgos crediticios. Estas aplicaciones evidencian cómo el aprendizaje automático puede revolucionar procesos y optimizar la toma de decisiones en múltiples industrias (Iberdrola, s.f.).

Pese a los beneficios que ofrece, el ML se enfrenta a numerosos retos. La calidad de los datos resulta fundamental; la presencia de datos incompletos o sesgados puede dar lugar a modelos que no funcionan adecuadamente. Asimismo, la capacidad de interpretación de los modelos genera debate, sobre todo en áreas críticas como la medicina o el sistema judicial, donde es vital comprender el proceso detrás de las decisiones tomadas. La exigencia de contar con un gran volumen de datos para entrenar modelos eficaces puede representar un obstáculo para varias organizaciones (Simeone, 2018).

2.26. Lenguaje de programación Python

Python se considera un lenguaje de programación de alto nivel que es interpretado y tiene un propósito general. Su desarrollo fue llevado a cabo por Guido van Rossum a finales de la década de 1980. Este lenguaje se diseñó con un enfoque en la claridad del código, lo que permite a los programadores transmitir ideas en menos líneas en comparación con otros lenguajes de programación. Se reconoce a Python por su facilidad de uso y su orientación hacia la productividad, lo que lo hace atractivo tanto para quienes están iniciando como para programadores con experiencia. Su sintaxis sencilla y el hecho de utilizar palabras clave en inglés, en lugar de símbolos complicados, hacen que el aprendizaje y la comprensión del lenguaje sean más accesibles.

La fama de Python ha experimentado un crecimiento notable en la última década, estableciéndolo como uno de los lenguajes de programación más empleados a nivel mundial. De acuerdo con el índice TIOBE, Python se sitúa actualmente en el tercer puesto entre los lenguajes más populares, solo superado por Java y C. Esta tendencia se atribuye a su versatilidad y a la abundancia de bibliotecas y frameworks disponibles, como Django para el desarrollo web y TensorFlow para la inteligencia artificial. Compañías líderes en

el sector tecnológico, como Google, Facebook y Netflix, incorporan Python en sus operaciones, lo que subraya su importancia en la industria (Gemrain, 2023). Python ha encontrado una amplia aceptación en el ámbito de la ciencia de datos y el ML. Su sintaxis sencilla y su extensa colección de funciones matemáticas y estadísticas lo posicionan como una herramienta excepcional para el análisis de datos y la creación de modelos predictivos. Bibliotecas como NumPy, Pandas y Matplotlib han convertido a Python en una elección habitual para la manipulación y visualización de datos (Gemrain, 2023).

2.27. Marco Legal

2.27.1. Estrategia de Gestión Integrada (EGI) de la OPS/OMS

Es un enfoque metódico para abordar la prevención y el control de enfermedades, en particular aquellas que son transmitidas por vectores, como el dengue, chikungunya y Zika. Se basa en la necesidad de llevar a cabo acciones coordinadas y de carácter multisectorial que incluyan a la comunidad, las autoridades sanitarias y otros sectores importantes, cuyo principal propósito es disminuir la morbilidad y mortalidad vinculadas a estas enfermedades mediante la promoción de la salud y la modificación de conductas de riesgo dentro de la población (OPS, 2023).

Desde su puesta en marcha, la Estrategia de Gestión Integrada (EGI) ha facilitado la consolidación de diversos elementos de la salud pública, abarcando la vigilancia epidemiológica, el manejo de casos y el control de vectores, que se centra en reaccionar ante brotes de enfermedades, y en buscar implementar medidas preventivas sostenibles que fortalezcan la capacidad de las comunidades para gestionar su propia salud. La EGI se fundamenta en la participación activa de la comunidad, promoviendo la educación y la sensibilización sobre los riesgos relacionados con las enfermedades arbovirales y la relevancia de la prevención (OPS, 2023).

La Estrategia de Gestión Integrada (EGI) también ha sido ajustada a las realidades locales de los países de la región, lo que permite que cada nación lleve a cabo acciones específicas fundamentadas en su contexto epidemiológico y social, abarcando la capacitación del personal de salud, la mejora de la infraestructura sanitaria y la promoción de políticas públicas que respalden la gestión integrada de la salud. Igualmente, la EGI subraya la relevancia de la investigación y la recopilación de datos para evaluar la efectividad de las intervenciones y adaptar las estrategias según sea necesario (OPS, 2023). Conforme a

Guzmán y Kourí (2002), la Estrategia de Gestión Integrada (EGI) ha mostrado ser una estrategia eficaz para el control de enfermedades arbovirales en la región de las Américas. Los autores subrayan que la consolidación de diversos elementos de la salud pública, tales como la vigilancia epidemiológica y el control de vectores, ha facilitado una respuesta más efectiva y coordinada ante los brotes de estas enfermedades. La participación de la comunidad y la educación sanitaria han sido esenciales para fomentar cambios en el comportamiento y disminuir los factores de riesgo relacionados con la transmisión de los virus.

Brathwaite Dick et al. (2012) subrayan la relevancia de ajustar la Estrategia de Gestión Integrada (EGI) a las realidades locales de cada nación. Los autores indican que el éxito de esta estrategia está ligado a la capacidad de los sistemas de salud para llevar a cabo intervenciones específicas de acuerdo con su contexto epidemiológico y social, lo que conlleva la necesidad de fortalecer las habilidades del personal de salud, mejorar la infraestructura sanitaria y fomentar políticas públicas que respalden la gestión integrada de la salud.

2.27.2. Reglamento Sanitario Internacional (RSI) 2005

Es un instrumento legal de la Organización Mundial de la Salud (OMS) cuyo propósito es prevenir, proteger, controlar y proporcionar una respuesta de salud pública adecuada y limitada a los riesgos que pueden extenderse a nivel internacional. Al mismo tiempo, busca evitar interferencias innecesarias con el tráfico y el comercio internacional. Fue adoptado en 2005 durante la 58ª Asamblea Mundial de la Salud y comenzó a ser aplicable en 2007. En la actualidad, es de cumplimiento obligatorio en 196 naciones, que incluyen a los 194 Estados Miembros de la OMS. El RSI exige que los países designen un punto focal nacional para las comunicaciones con la OMS y que establezcan capacidades esenciales para la vigilancia y la respuesta (OMS, 2005).

Una de las principales innovaciones del Reglamento Sanitario Internacional (RSI) de 2005 es la incorporación del concepto de "emergencia de salud pública de importancia internacional" (ESPII). Se entiende por ESPII a un evento excepcional que representa un riesgo para la salud pública de otros países debido a la propagación internacional de una enfermedad y que podría requerir una respuesta coordinada a nivel internacional (Rodier et al., 2007). El RSI (2005) también establece importantes medidas para salvaguardar los

derechos de los viajeros y otras personas, abordando aspectos como el tratamiento de datos personales, el consentimiento informado y la no discriminación en la aplicación de medidas sanitarias. Adicionalmente, fomenta la cooperación internacional y el intercambio de información con el fin de mejorar la preparación y la respuesta ante eventos que pudieran ser considerados como una ESPII (Rodier et al., 2007).

2.273. Directrices de la Organización Mundial de la Salud (OMS)

Son herramientas esenciales que ofrecen recomendaciones fundamentadas en evidencia para mejorar la salud pública y orientar a los países en la implementación de políticas efectivas. Estas directrices abarcan una amplia variedad de temas, desde el tratamiento de enfermedades hasta la promoción de estilos de vida saludables. Su finalidad es asistir a los gobiernos, a los profesionales de la salud y a otros intervinientes en la toma de decisiones informadas que contribuyan a mejorar la salud de las poblaciones (World Health Organization [WHO], 2021). Las directrices de la OMS ofrecen datos estandarizados y fiables sobre la epidemiología del dengue, incluyen mejores prácticas y protocolos para la vigilancia y el control del dengue, lo que facilita el diseño de modelos que reflejan escenarios realistas y eficaces (OMS, 2023).

La Organización Mundial de la Salud (OMS, 2023) fomenta la integración de datos que abarcan aspectos epidemiológicos, clínicos, laboratoriales y entomológicos, lo que posibilita una simulación más exhaustiva y precisa. Las directrices de la OMS son actualizadas de manera regular para reflejar los últimos avances en ciencia y tecnología, ofreciendo un marco que es aplicable tanto a nivel global como local para la gestión del dengue, lo cual permite que los modelos sean flexibles, adaptándose a diversos contextos geográficos y socioeconómicos.

2.274. Programa Nacional de Control Vectorial del Dengue. Servicio Nacional de Erradicación del Paludismo

Llevado a cabo por el Servicio Nacional de Erradicación del Paludismo (SENEPA) en Paraguay, representa una estrategia integral diseñada para disminuir la incidencia del dengue y de otras enfermedades transmitidas por vectores, como el Zika y el chikungunya. El programa se fundamenta en un enfoque de vigilancia y control continuo del mosquito *Aedes aegypti*, que es el principal vector de estas enfermedades. Las acciones se realizan a lo largo de todo el año, adaptándose a las variaciones estacionales y epidemiológicas que impactan la proliferación de los vectores (SENEPA, 2023).

El programa incluye diversos elementos esenciales, tales como el saneamiento ambiental, la promoción de la salud y una comunicación efectiva con la comunidad, que tienen como objetivo no solo controlar la población de mosquitos, sino también educar a la ciudadanía sobre la importancia de eliminar los criaderos y adoptar prácticas de prevención. La participación comunitaria es clave en este enfoque, ya que se busca empoderar a los ciudadanos para que tomen medidas proactivas en la lucha contra el dengue (Ministerio de Salud Pública y Bienestar Social [MSPBS], 2023). Pese a los esfuerzos realizados, el dengue continúa siendo un desafío considerable para la salud pública en Paraguay. Las condiciones climáticas, sumadas a la urbanización y a los cambios en los patrones de comportamiento de la población, juegan un papel importante en la persistencia del dengue en el país. Por lo tanto, resulta fundamental seguir fortaleciendo las capacidades del SENEPA y promover la educación y la participación de la comunidad para alcanzar un control sostenible de las enfermedades transmitidas por vectores (SENEPA, 2023).

2.275. Política Nacional de Salud 2015-2030

Tiene como meta primordial asegurar el acceso y la cobertura universal de salud en todo el territorio, en alineación con el derecho a la salud que se encuentra en la Constitución Nacional y los compromisos internacionales asumidos por el país. Se organiza en torno a la mejora de la equidad, la calidad de los servicios de salud y la disminución de las desigualdades sanitarias (MSPBS, 2016).

Según lo indicado por el MSPBS (2016), el documento fue ratificado a través del Decreto N° 4541/2015 y la Resolución S.G. N° 612, lo que lo establece como una política de Estado, sin importar los cambios de gobierno, encontrándose en consonancia con la estrategia de la OPS para avanzar hacia el acceso y la cobertura universal de salud, eliminando las barreras económicas, geográficas y culturales que impiden el acceso a los servicios de salud. Los objetivos de esta política son disminuir las inequidades en salud y mejorar la calidad de vida de la población mediante el fortalecimiento del sistema de salud, promoviendo tanto el acceso universal como la cobertura integral. Se centra en la atención primaria como el eje fundamental para alcanzar estos objetivos, reconociendo que un sistema de salud más accesible y equitativo es vital para enfrentar las disparidades en la atención.

3. Marco Metodológico

En este apartado se expone el enfoque metodológico de investigación, el diseño de investigación y el tipo de investigación. Además, se describe las etapas de la metodología para investigaciones tecnológicas y científicas según Cegarra Sánchez (2004).

3.1. Enfoque de Investigación

El enfoque de este trabajo es mixto, ya que combina tanto métodos cualitativos como cuantitativos para mejorar la comprensión del comportamiento del dengue y generar predicciones precisas. Este enfoque permite analizar diferentes fuentes de datos, como los datos epidemiológicos y meteorológicos, lo que contribuye a un análisis integral del problema (Cresswell, 2003). Desde el **método cualitativo**, el modelo de ML buscó ser evaluado críticamente en sus predicciones de dengue, a través del criterio experto del investigador. Gracias al uso de datos históricos y el presente conocimiento subyacente de los patrones de transmisión del dengue, el modelo fue ajustado de manera iterativa, permitiendo al investigador afinar los parámetros para la mejora progresiva de su capacidad predictiva en tiempo real. A su vez, los datos meteorológicos como la temperatura y la humedad fueron añadidos para entender de mejor manera las condiciones de transmisión del dengue, lo cual posibilitó al investigador contextualizar las proyecciones hechas, ubicando geográficamente los riesgos de acuerdo a dichas condiciones climáticas, mejorando así la capacidad del modelo de otorgar predicciones más ajustadas a la realidad.

En lo que refiere al **método cuantitativo**, se utilizó un modelo de red neuronal LSTM para trabajar con series temporales y anticipar futuros brotes de dengue, lo que posibilitó evaluar precisamente las predicciones del modelo a través del uso de métricas objetivas. Específicamente, el modelo se ajustó y evaluó en función de su precisión, comparando sus proyecciones con valores reales. Se procesaron registros de movimientos epidemiológicos y meteorológicos, generando pronósticos diarios, que fueron validados y comparados con datos históricos sobre patrones de propagación conocidos. Durante el desarrollo, el rendimiento del modelo se midió mediante el uso de métricas como el *Mean Squared Error* (MSE) y el coeficiente de correlación. De este modo, se ajustaron iterativamente las proyecciones para optimizar la capacidad predictiva del modelo,

asegurando que las estimaciones muestren con precisión el comportamiento dinámico actual del dengue.

3.2. Diseño de Investigación

El diseño de este trabajo final de grado fue exploratorio, como lo indica Cresswell (2003), en la caracterización de investigaciones con enfoque mixto. La finalidad de este tipo de estudio es examinar un tema o problema de investigación poco estudiado o que no ha sido abordado previamente, y construir una comprensión basada en los resultados obtenidos. Atendiendo a lo mencionado, esta investigación fue de carácter exploratorio, debido a que se trató de una evaluación y prueba de funcionalidad de un sistema innovador de monitoreo epidemiológico en tiempo real, denominado “Simulación dinámica en tiempo real del comportamiento de la propagación del dengue a través de ML”. La novedad de este enfoque consistió en la implementación de una herramienta que combina datos meteorológicos y epidemiológicos en tiempo real, para modelar el crecimiento y propagación del dengue. La herramienta, diseñada para el uso en el Hospital de Santa Rosa Misiones, posibilitó predecir de forma temprana picos de contagio, para la planificación de estrategias de respuesta y prevención.

Al ser una simulación en tiempo real, se evaluaron maneras innovadoras de implementar el uso de herramientas de ML con datos dinámicos y contextuales. Debido a que este enfoque no se ha llevado a cabo previamente en el entorno de la localidad mencionada ni con las características de una simulación continua, la presente investigación se considera exploratoria al ahondar en la precisión, eficacia y viabilidad de dicho sistema en un contexto cambiante y práctico.

3.3. Tipos de la investigación

La investigación es de tipo aplicada, ya que se enfoca en utilizar y adaptar conocimientos científicos derivados de investigaciones previas (Cegarra Sánchez, 2004). El propósito de esta investigación fue llevar a cabo una solución práctica, innovadora y confiable para abordar un problema específico sobre la vigilancia y predicción de la propagación del dengue. La investigación se clasifica como aplicada, debido a que se combinan las herramientas de software y tecnologías ya existentes para la creación de un sistema que

puede ser usado en entornos epidemiológicos y hospitalarios. Dicha solución constituye una mejora tangible en el aspecto de vigilancia y control de brotes y ha sido diseñada para ser ubicua y accesible a los centros pertinentes en términos de implementación y vigilancia de posibles picos de contagio.

3.4. Etapas Metodológicas

Para llevar a cabo los objetivos planteados en este trabajo final de grado se siguieron las siguientes etapas, las cuales estuvieron basadas en la metodología de investigación científica y tecnológica desarrollada en Cegarra Sánchez (2004).

3.4.1. Etapa 1: Exploratoria

En esta etapa inicial, se llevó a cabo una revisión sistemática y extensa de la literatura relacionada con las herramientas, algoritmos, modelos de ML y software disponibles para la simulación de la propagación de enfermedades como el dengue, incluyéndose estudios sobre la relación entre los factores epidemiológicos y meteorológicos, así como las tecnologías aplicadas para la simulación en tiempo real. A partir de los requisitos del proyecto, se eligieron las herramientas y la metodología que se consideraron más adecuadas para el desarrollo de una simulación destinada a predecir el comportamiento de la propagación del dengue y la detección de la ubicación de futuros brotes.

3.4.1.1. Visita al Hospital Santa Rosa Misiones

En esta etapa se realizó la visita al Hospital Santa Rosa de Misiones con el fin de recopilar los datos epidemiológicos y meteorológicos necesarios para la simulación dinámica en tiempo real del comportamiento de la propagación de dengue a partir de técnicas de ML. A continuación, se desglosa la información utilizada como base de datos para el desarrollo del modelo predictivo:

- **Fecha de los primeros síntomas:** Constituye un dato fundamental para precisar el inicio de la propagación del dengue, facilitando la identificación de patrones temporales clave que posibilitan anticipar futuros brotes y analizar la evolución de los casos en el tiempo.
- **Ubicación (barrio o compañía):** Posibilita que la simulación identifique las áreas específicas donde se han detectado casos de dengue, lo cual contribuye a

ajustar el modelo y perfeccionar la precisión de las predicciones.

- **Edad y género de los pacientes:** Son esenciales para identificar los grupos poblacionales más vulnerables al dengue. La segmentación de los pacientes por edad y género permite el ajuste del modelo, mejorando la precisión de las predicciones sobre los grupos de riesgo.
- **Temperatura máxima y mínima:** Las condiciones climáticas, especialmente las variaciones de temperatura, influyen directamente en la supervivencia y capacidad reproductiva de los mosquitos *Aedes aegypti*.
- **Viento (km/h) y dirección del viento:** Con estos datos se comprende el efecto del viento en la dispersión de los mosquitos, puesto que la velocidad y dirección del mismo pueden influir en la propagación del dengue al afectar el desplazamiento de los vectores en áreas adyacentes.
- **Precipitación:** Debido a que el *Aedes aegypti* deposita sus huevos en agua estancada, la cantidad de lluvia es un factor clave en la formación de criaderos de mosquitos; en consecuencia, un incremento en las precipitaciones puede correlacionarse con un aumento en los casos de dengue.

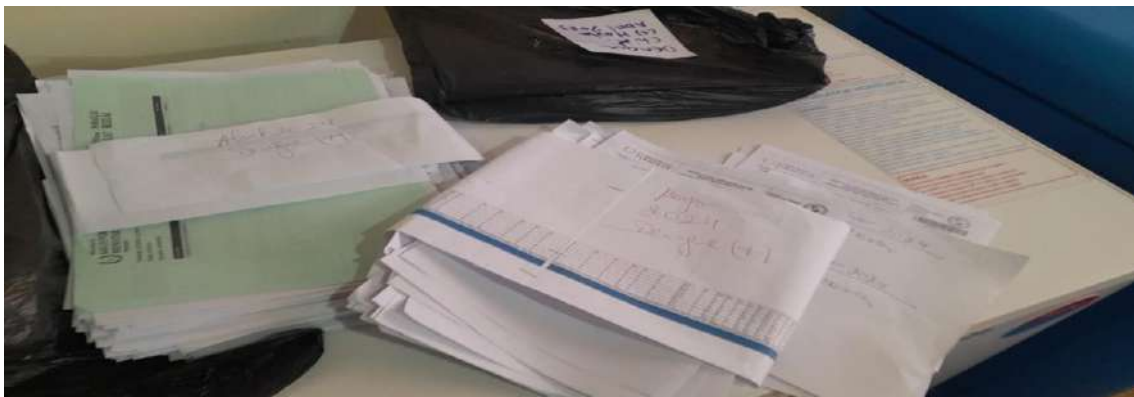


Figura Nro 1 Recolección de datos epidemiológicos 2019- 2024

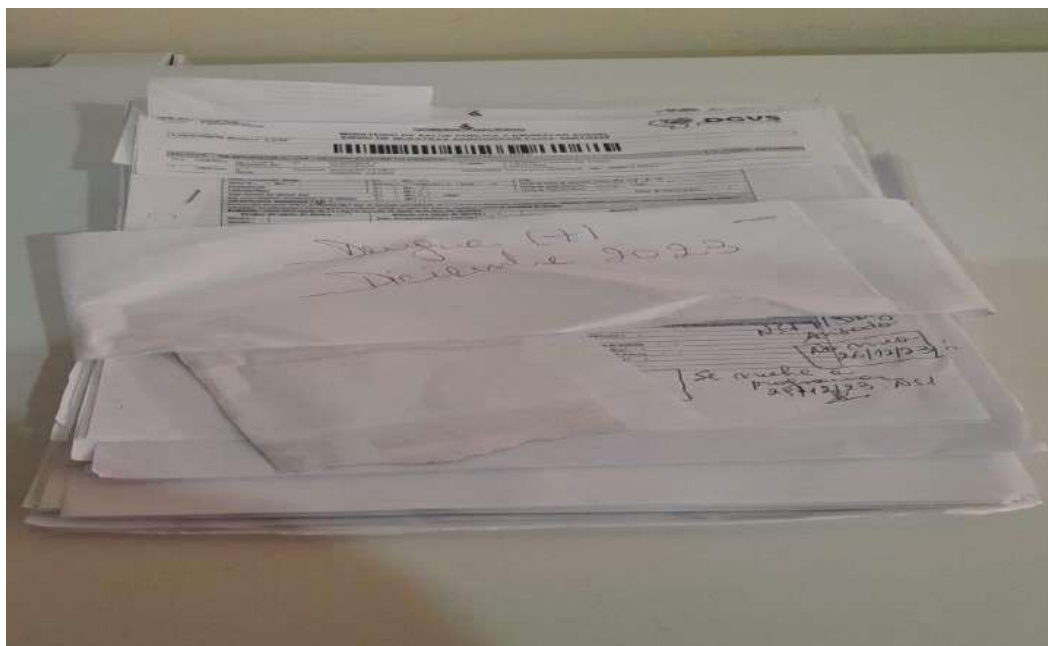


Figura Nro 2 Archivo de datos epidemiológicos diciembre 2023

Universidad Nacional de Asunción
FACULTAD DE CIENCIAS AGRARIA – FILIAL SANTA ROSA MNES.

PLANILLA DE DATOS METEOROLÓGICOS

Mes: MARZO Año: 2024

Fecha	Hora	T° seco	T° hum.	H° relativa	T° max.	T° mín.	Viento Km/H	Direc. Viento	Est. Del tiempo	Evapo ración	Precipit acción	Pres. atmosf	Nivel agua
1					22	19	1	S/E	P/N	-	-	991	
					28	26	1	S/E	P/N	-	-	991	
Resumen													
2					23	19	2	S/E	P/N	-	-	991	
					31	29	1	S/O	D	-	-	989	
Resumen													
3					29	22	25	E	D.	-	-	989	
					23	21	2	S/E	D	-	15	991	
Resumen													
4					22	20	7	S/E	N	-	-	990	
					26	22	1	S/E	P	-	5mm	990	
Resumen													
5					23	19	14	S/E	N	-	-	991	
					26	24	10	S	D	-	-	992	
Resumen													

Figura Nro 3 Planilla de datos meteorológicos marzo 2024

Fecha	Hora	T° seco	T° hum.	H° relativa	T° max.	T° min.	Viento Km/H	Direc. Viento	Est. Del tiempo	Evapo ración	Precipit acción	Pres. atmosf	Nivel agua
16					29	18	1	S/O	N	-	-	992	
					26	24	2	S/O	D	-	-	992	
Resumen													
17					21	19	2	S	D	-	-	992	
					28	22	3	S/O	D	-	-	990	
Resumen													
18					21	17	1	S/O	D	-	-	990	
					29	23	1	S	P/N	-	-	989	
Resumen													
19					24	19	2	S/O	P/N	-	-	989	
					25	23	5	S/E	P/N	-	-	989	
Resumen													
20					22	20	10	S/O	D	-	-	989	
					28	22	7	S/E	P/N	-	-	988	
Resumen													
21					23	19	5	S/E	D	-	-	989	
					29	24	2	S/E	D	-	-	988	
Resumen													
22					23	19	1	S/E	P/N	-	-	988	

Figura Nro 4 Planilla de datos meteorológicos 2019 – 2023

Fecha	Hora	T° seco	T° hum.	H° relativa	T° max.	T° min.	Viento Km/H	Direc. Viento	Est. Del tiempo	Evapo ración	Precipit acción	Pres. atmosf	Nivel agua
1													
Resumen													
2													
					27	20	3	S/O	P/N	-	-	990	
Resumen													
3					22	18	5	S/O	D	-	-	990	
					28	24	1	S/E	D	-	-	990	
Resumen													
4					23	19	3	S/O	D	-	-	990	
					28	22	1	S/O	P/N	-	-	991	
Resumen													
5					23	19	1	S	D	-	-	991	
					29	23	2	S/E	P/N	-	-	990	
Resumen													
6					22	19	2	S	P/N	-	-	990	
					30	27	1	N/E	P/N	-	-	989	
Resumen													
7					24	20	3	N/E	P/N	-	-	989	
					30	25	2	S/E	P/N	-	-	989	
Resumen													

Figura Nro 5 Planilla de datos meteorológicos abril 2021-2023

Estación Meteorológica: F.C.A. - U.N.A.								
Santa Rosa - Misiones Enero 2022.								
Fecha	Hora	T° max °C	T° min °C	Precipitación (mm)	T° Max del día	T° Min del día	T° Media diaria	Total Precipitación (mm)
1	08:00	27	24.0		34	27	30.5	0
	15:00	39	34.0					
	20:00	29	24.0					
2	08:00	27	22.0		30	26	28.0	0
	15:00	32	30.0					
	20:00	26	24.0					
3	08:00	26	22.0		27.0	26	26.5	0
	15:00	31	27.0					
	20:00	27	24.0					
4	08:00	25	22.0		22.0	25	23.5	0
	15:00	35	22.0					
	20:00	31	22.0					
5	08:00	26	19.0		31	26	28.5	0
	15:00	33	31.0					
	20:00	26	24.0					
6	08:00	24	20.0		27	24	25.5	0
	15:00	33	27.0					
◀▶ Enero 2022 Febrero 2022 Marzo 2022 Abril 2022 Mayo 2022								

Figura Nro 6 Datos meteorológicos – año 2022

3.4.1.2. Inserción de datos en Excel

Datos Epidemiológicos 2019 Santa Rosa Misiones				Datos Epidemiológicos 2019 Santa Rosa Misiones			
Febrero 2019		Barrio: Pablo VI		Febrero 2020		Compañía: Distintas compañía	
Fecha	Ubicación	Edad	Genero	Fecha	Ubicación	Edad	Genero
05/02/2019	Pablo VI	36	M	21/02/2020	Colonia Acevedo	10	F
01/02/2019	Pablo VI	70	M	20/02/2020	Colonia Acevedo	11	M
02/02/2019	Pablo VI	56	F	23/02/2020	3 de Mayo	57	F
03/02/2019	Pablo VI	22	M	27/02/2020	San Gabriel	30	F
04/02/2019	Pablo VI	47	M	22/02/2020	San Gabriel	15	F
05/02/2019	Pablo VI	30	M	22/02/2020	San Francisco	22	F
06/02/2019	Pablo VI	78	M	21/02/2020	Colonia Acevedo	11	M
07/02/2019	Pablo VI	15	M	22/02/2020	San Francisco	35	m
08/02/2019	Pablo VI	7	M	19/02/2020	Gabino Roja	67	F
09/02/2019	Pablo VI	73	M	14/02/2024	San Francisco	53	F
10/02/2019	Pablo VI	12	F	15/02/2020	Cerro Costa	14	M
11/02/2019	Pablo VI	44	F	10/02/2020	San Jose	83	M
12/02/2019	Pablo VI	14	M	14/02/2020	Gabino Roja	30	F
13/02/2019	Pablo VI	34	M	12/02/2020	San Francisco	4	M
14/02/2019	Pablo VI	66	F	14/02/2020	San Solano	20	M
15/02/2019	Pablo VI	81	F	07/02/2020	San Jose	20	M
16/02/2019	Pablo VI	66	F	08/02/2020	Jacarey	17	F
17/02/2019	Pablo VI	8	M	07/02/2020	Colonia Acevedo	66	M

Figura Nro 7 Datos cargados en Excel

Febrero 2021 Barrio: San Jose y Cristo Rey			
Fecha	Ubicación	Edad	Genero
16/02/2021	Sa Jose	67	F
11/02/2021	Cristo Rey	55	F
14/02/2021	Cristo Rey	35	F
14/02/2021	Sa Jose	28	F
13/02/2021	Sa Jose	29	M
17/02/2021	Cristo Rey	59	F
21/02/2021	Sa Jose	43	M
27/02/2021	Sa Jose	18	M
23/02/2021	Sa Jose	9	M
26/02/2021	Sa Jose	79	F
23/02/2021	Sa Jose	20	M
22/02/2021	Sa Jose	36	F
14/02/2021	Cristo Rey	81	F
19/02/2021	Cristo Rey	34	F
13/02/2021	Sa Jose	41	F
13/02/2021	Sa Jose	40	F
11/02/2021	Cristo Rey	38	M
12/02/2021	Sa Jose	44	F
11/02/2021	Sa Jose	46	F
09/02/2021	Cristo Rey	56	M
10/02/2021	Cristo Rey	25	F
06/02/2021	Cristo Rey	34	M

Figura Nro 8 Datos cargados en Excel (continuación)

Datos Epidemiológicos 2022 Santa Rosa Misiones				Datos Epidemiológicos 2023 Santa Rosa Misiones			
Febrero 2022 Barrio: San Isidro				Febrero 2023 Barrio: Virgen del pilar			
Fecha	Ubicación	Edad	Genero	Fecha	Ubicación	Edad	Genero
28/03/2022	San Isidro	26	F	12/03/2023	Virgen del pilar	43	F
29/03/2022	San Isidro	4	M	11/03/2023	Virgen del pilar	43	M
30/03/2022	San Isidro	2	F	26/03/2023	Virgen del pilar	55	M
31/03/2022	San Isidro	16	M	25/03/2023	Virgen del pilar	7	F
01/04/2022	San Isidro	29	M	25/03/2023	Virgen del pilar	73	M
02/04/2022	San Isidro	6	F	24/03/2023	Virgen del pilar	24	F
03/04/2022	San Isidro	29	M	22/03/2023	Virgen del pilar	12	M
04/04/2022	San Isidro	24	F	20/03/2023	Virgen del pilar	78	F
05/04/2022	San Isidro	16	F	23/03/2023	Virgen del pilar	19	F
06/04/2022	San Isidro	18	F	23/03/2023	Virgen del pilar	15	F
07/04/2022	San Isidro	41	F	23/03/2023	Virgen del pilar	10	M
08/04/2022	San Isidro	6	F	25/03/2023	Virgen del pilar	75	F
09/04/2022	San Isidro	28	F	25/03/2023	Virgen del pilar	29	F
10/04/2022	San Isidro	13	F	24/03/2023	Virgen del pilar	33	F
11/04/2022	San Isidro	80	M	22/03/2023	Virgen del pilar	45	M
12/04/2022	San Isidro	41	F	20/03/2023	Virgen del pilar	6	M
13/04/2022	San Isidro	3	F	23/03/2023	Virgen del pilar	67	M
14/04/2022	San Isidro	30	M	23/03/2023	Virgen del pilar	30	F

Figura Nro 9 Datos cargados en Excel (continuación)

Febrero 2023 Barrio: San Miguel				Febrero 2022 Barrio: San Jose			
Fecha	Ubicación	Edad	Genero	Fecha	Ubicación	Edad	Genero
12/03/2023	San Miguel	50	F	28/03/2022	San Jose	9	M
11/03/2023	San Miguel	61	M	29/03/2022	San Jose	20	F
26/03/2023	San Miguel	4	M	30/03/2022	San Jose	82	F
25/03/2023	San Miguel	58	F	31/03/2022	San Jose	86	F
25/03/2023	San Miguel	55	M	01/04/2022	San Jose	49	M
24/03/2023	San Miguel	51	F	02/04/2022	San Jose	57	F
22/03/2023	San Miguel	50	F	03/04/2022	San Jose	29	M
20/03/2023	San Miguel	17	M	04/04/2022	San Jose	39	F
23/03/2023	San Miguel	38	M	05/04/2022	San Jose	7	M
23/03/2023	San Miguel	33	F	06/04/2022	San Jose	37	F
23/03/2023	San Miguel	30	F	07/04/2022	San Jose	25	F
29/03/2023	San Miguel	20	M	08/04/2022	San Jose	16	F
25/03/2023	San Miguel	29	M	09/04/2022	San Jose	36	M
22/03/2023	San Miguel	56	F	10/04/2022	San Jose	25	F
11/03/2023	San Miguel	40	M	11/04/2022	San Jose	40	F
26/03/2023	San Miguel	30	F	12/04/2022	San Jose	49	F
25/03/2023	San Miguel	9	F	13/04/2022	San Jose	15	M
25/03/2023	San Miguel	74	F	14/04/2022	San Jose	57	F

Figura Nro 10 Datos epidemiológicos – 2022-2023

Febrero 2022 Barrio: Distintas compañías			
Fecha	Ubicación	Edad	Genero
28/03/2022	san jose ita moroti	1	F
29/03/2022	san jose ita moroti	28	F
30/03/2022	San Francisco	13	M
31/03/2022	San Francisco	69	F
01/04/2022	San Francisco	2	M
02/04/2022	San Francisco	15	M
03/04/2022	San Francisco	41	F
04/04/2022	San Solano	58	F
05/04/2022	Itapypore	29	F
06/04/2022	Potrero alto	41	F
07/04/2022	Potrero alto	35	F
08/04/2022	Potrero alto	15	F
09/04/2022	Ykua sati	52	F
10/04/2022	San Francisco	66	F
11/04/2022	San Francisco	54	M
12/04/2022	Ykua sati	8	F
13/04/2022	Ykua sati	29	M
14/04/2022	Gabino Roja	6	F

Figura Nro 11 Datos epidemiológicos 2022

Así, se procedió a organizar la información en Excel de la forma más metódica posible, por lo que todas las variables fueron estructuradas y etiquetadas. Para la organización se tuvo en cuenta la clasificación de la información por tipo, tales como los registros de casos de dengue, temperatura, humedad y otras variables importantes. De esta manera, es posible su fácil entendimiento y manejo. Como se puede apreciar en las imágenes del proceso, la información fue dividida en años, pudiendo ver de manera detallada y desglosada los datos epidemiológicos y meteorológicos correspondientes que pudieron influir en el mismo.

Colocar la información en Excel de manera ordenada ayuda a ver las tendencias y a identificar patrones anuales, siendo a la vez un prelude para la próxima fase del proyecto, que es cargar los datos en la base de datos. De esta manera, al ser ordenados según el año, estos se vuelven más fáciles de combinar y cargar, garantizando así que la base de datos esté bien estructurada. La estructura mencionada es importante para garantizar que el sistema de ML pueda recuperar la información rápidamente y con precisión, lo que es esencial para la predicción y para la posterior comparación con los hechos reales de propagación de dengue.

Datos Epidemiológicos 2023 Santa Rosa Misiones			
Febrero 2023 Barrio: Virgen del pilar			
Fecha	Ubicación	Edad	Genero
12/03/2023	Virgen del pilar	43	F
11/03/2023	Virgen del pilar	43	M
26/03/2023	Virgen del pilar	55	M
25/03/2023	Virgen del pilar	7	F
25/03/2023	Virgen del pilar	73	M
24/03/2023	Virgen del pilar	24	F
22/03/2023	Virgen del pilar	12	M
20/03/2023	Virgen del pilar	78	F
23/03/2023	Virgen del pilar	19	F
23/03/2023	Virgen del pilar	15	F
23/03/2023	Virgen del pilar	10	M
25/03/2023	Virgen del pilar	75	F
25/03/2023	Virgen del pilar	29	F
24/03/2023	Virgen del pilar	33	F
22/03/2023	Virgen del pilar	45	M
20/03/2023	Virgen del pilar	6	M
23/03/2023	Virgen del pilar	67	M
23/03/2023	Virgen del pilar	30	F

Figura Nro 12 *Preparación de datos para cargar en la base de datos*

Fecha	Ubicación	Edad	Genero
05/02/2019	Pablo VI	36	M
01/02/2019	Pablo VI	70	M
02/02/2019	Pablo VI	56	F
03/02/2019	Pablo VI	22	M
04/02/2019	Pablo VI	47	M
05/02/2019	Pablo VI	30	M
06/02/2019	Pablo VI	78	M
07/02/2019	Pablo VI	15	M
08/02/2019	Pablo VI	7	M
09/02/2019	Pablo VI	73	M
10/02/2019	Pablo VI	12	F
11/02/2019	Pablo VI	44	F
12/02/2019	Pablo VI	14	M
13/02/2019	Pablo VI	34	M
14/02/2019	Pablo VI	66	F
15/02/2019	Pablo VI	81	F
16/02/2019	Pablo VI	66	F

Figura Nro 13 Preparación de datos para cargar en la base de datos (continuación)

3.4.13. Carga de datos históricos en la base de datos

```

datos_dengue.csv > data
Fecha,Ubicación,Edad,Genero
01/01/2019,Virgen del pilar,33,M # Inicio de datos epidemiológicos 2019
01/01/2019,San Jose,67,F
01/03/2019,virgen del pilar,33,F
01/03/2019,virgen del pilar,67,M
01/04/2019,Cerro Costa,20,F
01/04/2019,San Jose,19,M
01/05/2019,Pablo VI,64,F
01/06/2019,San Geronimo,27,M
01/07/2019,San Jose ita moroti,27,M
02/04/2019,San Geronimo,38,F
02/04/2019,San Geronimo,39,M
02/06/2019,Pablo VI,50,F
02/08/2019,Pablo VI,14,F
02/08/2019,San Geronimo,38,F
02/09/2019,Virgen del pilar,12,F
02/09/2019,Virgen del pilar,16,F
02/10/2019,San Miguel,27,F
03/03/2019,San Isidro,22,F
03/09/2019,Virgen del pilar,16,F
04/03/2019,San Geronimo,11,M
04/03/2019,san isidro,20,F
04/04/2019,San Geronimo,11,M
04/10/2019,San Miguel,25,F
05/03/2019,San Isidro,87,M
05/04/2019,San Isidro,87,M
05/04/2019,Yacarey,2,M
05/08/2019,Catetral,17,F
06/03/2019,San Isidro,19,M
06/03/2019,San isidro,24,M
06/03/2019,Virgen del pilar,1,M
06/04/2019,Cerro Costa,24,M

```

Figura Nro 14 Datos epidemiológicos 2019

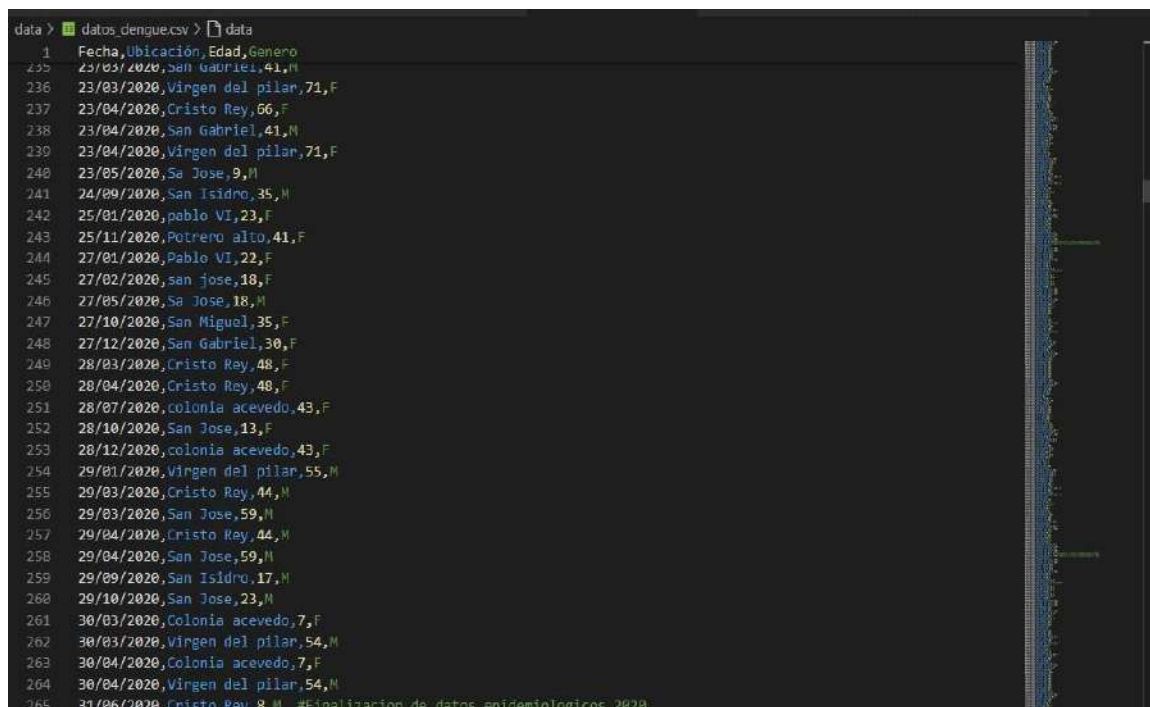
```
ata > datos.dengue.csv > data
```

	1	Fecha,Ubicación,Edad,Genero
104	23/03/2019,Virgen del pilar,71,F	
105	23/04/2019,Cristo Rey,66,F	
106	23/04/2019,San Gabriel,41,M	
107	23/04/2019,Virgen del pilar,71,F	
108	23/05/2019,Sa Jose,9,M	
109	24/09/2019,San Isidro,35,M	
110	25/01/2019,pablo VI,23,F	
111	25/11/2019,Potrero alto,41,F	
112	27/01/2019,Pablo VI,22,F	
113	27/02/2019,san jose,18,F	
114	27/05/2019,Sa Jose,18,M	
115	27/10/2019,San Miguel,35,F	
116	27/12/2019,San Gabriel,30,F	
117	28/03/2019,Cristo Rey,48,F	
118	28/04/2019,Cristo Rey,48,F	
119	28/07/2019,colonia acevedo,43,F	
120	28/10/2019,San Jose,13,F	
121	28/12/2019,colonia acevedo,43,F	
122	29/01/2019,Virgen del pilar,55,M	
123	29/03/2019,Cristo Rey,44,M	
124	29/03/2019,San Jose,59,M	
125	29/04/2019,Cristo Rey,44,M	
126	29/04/2019,San Jose,59,M	
127	29/09/2019,San Isidro,17,M	
128	29/10/2019,San Jose,23,M	
129	30/03/2019,Colonia acevedo,7,F	
130	30/03/2019,Virgen del pilar,54,M	
131	30/04/2019,Colonia acevedo,7,F	
132	30/04/2019,Virgen del pilar,54,M	
133	31/06/2019,Cristo Rey,8,M	#Finalización de datos epidemiológicos 2019

Figura Nro 15 Datos epidemiológicos 2019 (continuación)

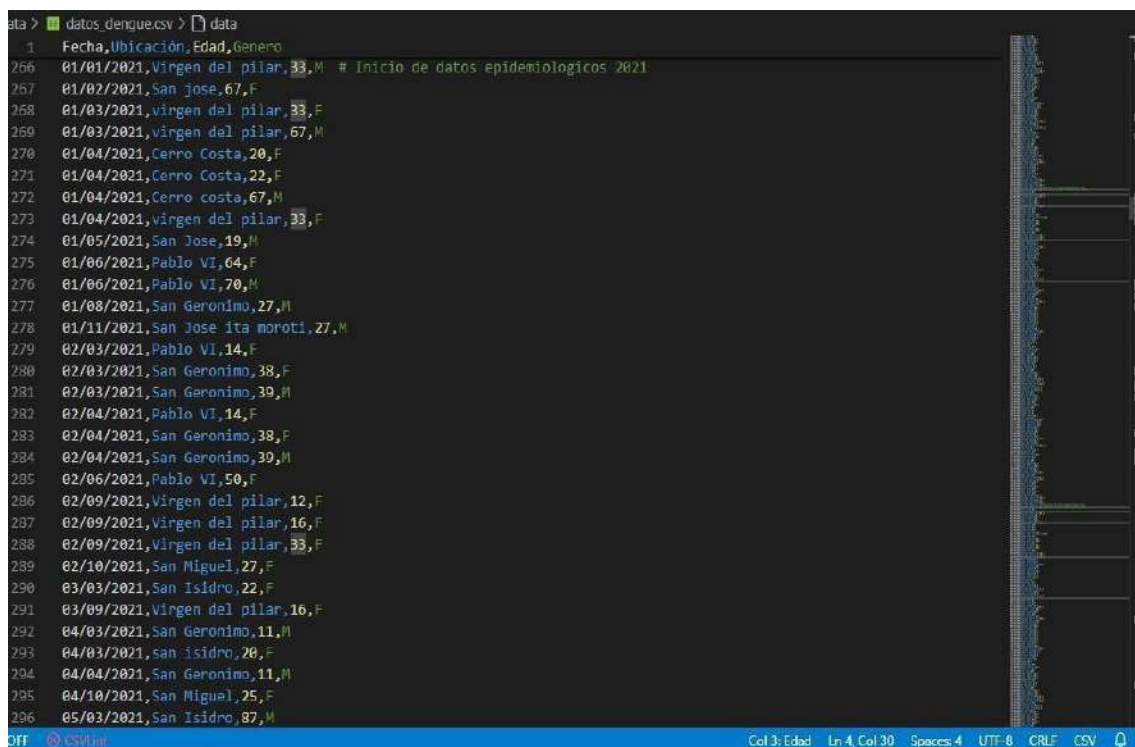
01/01/2020, Virgen del pilar, 33, M # Inicio de datos epidemiológicos 2020
01/02/2020, San jose, 67, F
01/03/2020, virgen del pilar, 33, F
01/03/2020, virgen del pilar, 67, M
01/04/2020, Cerro Costa, 20, F
01/05/2020, San Jose, 19, M
01/06/2020, Pablo VI, 64, F
01/07/2020, San Geronimo, 27, M
01/07/2020, San Jose ita moroti, 27, M
02/08/2020, Pablo VI, 14, F
02/08/2020, San Geronimo, 38, F
02/04/2020, San Geronimo, 38, F
02/04/2020, San Geronimo, 39, M
02/06/2020, Pablo VI, 50, F
02/09/2020, Virgen del pilar, 12, F
02/09/2020, Virgen del pilar, 16, F
02/10/2020, San Miguel, 27, F
03/03/2020, San Isidro, 22, F
03/09/2020, Virgen del pilar, 16, F
04/03/2020, San Geronimo, 11, M
04/03/2020, san isidro, 20, F
04/04/2020, San Geronimo, 11, M
04/10/2020, San Miguel, 25, F
05/03/2020, San Isidro, 87, M
05/04/2020, San Isidro, 87, M
05/04/2020, Vacarey, 2, M
05/08/2020, Catetral, 17, F
06/03/2020, San isidro, 19, M
06/03/2020, San isidro, 24, M
06/03/2020, Virgen del pilar, 1, M

Figura Nro 16 Datos epidemiológicos 2020



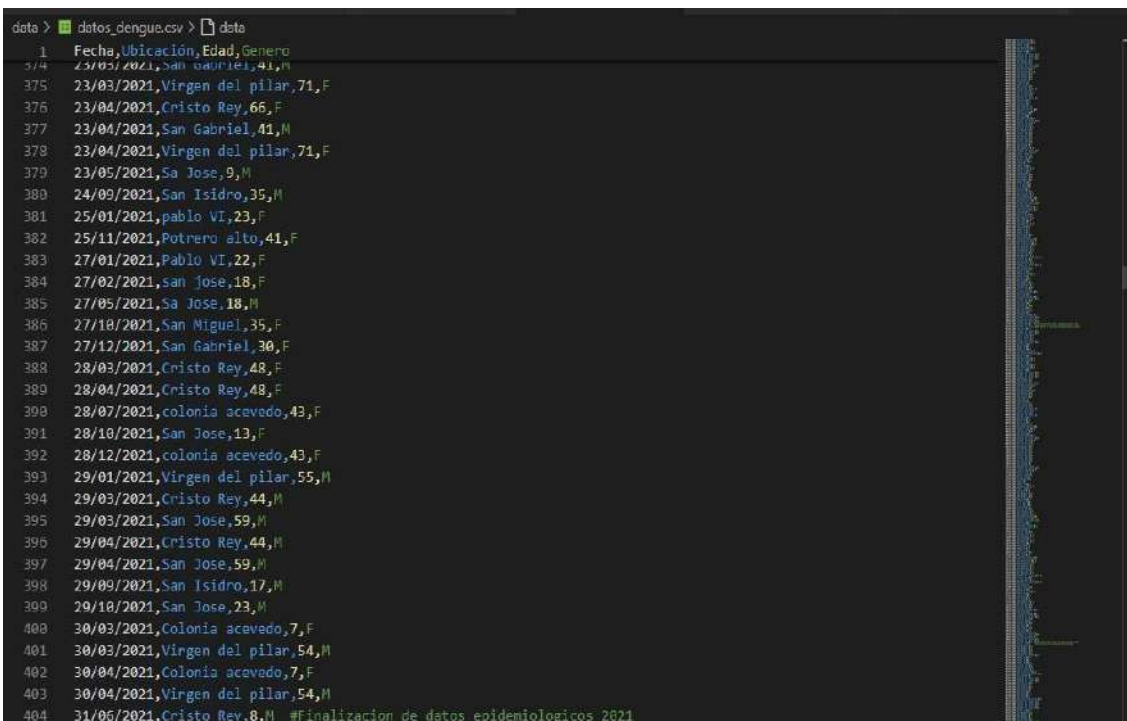
	Fecha	Ubicación	Edad	Género
1	23/03/2020	San Gabriel	41	M
236	23/03/2020	Virgen del pilar	71	F
237	23/04/2020	Cristo Rey	66	F
238	23/04/2020	San Gabriel	41	M
239	23/04/2020	Virgen del pilar	71	F
240	23/05/2020	San Jose	9	M
241	24/09/2020	San Isidro	35	M
242	25/01/2020	peblo VI	23	F
243	25/11/2020	Potrero alto	41	F
244	27/01/2020	Pablo VI	22	F
245	27/02/2020	san jose	18	F
246	27/05/2020	San Jose	18	M
247	27/10/2020	San Miguel	35	F
248	27/12/2020	San Gabriel	30	F
249	28/03/2020	Cristo Rey	48	F
250	28/04/2020	Cristo Rey	48	F
251	28/07/2020	colonia acevedo	43	F
252	28/10/2020	San Jose	13	F
253	28/12/2020	colonia acevedo	43	F
254	29/01/2020	Virgen del pilar	55	M
255	29/03/2020	Cristo Rey	44	M
256	29/03/2020	San Jose	59	M
257	29/04/2020	Cristo Rey	44	M
258	29/04/2020	San Jose	59	M
259	29/09/2020	San Isidro	17	M
260	29/10/2020	San Jose	23	M
261	30/03/2020	Colonia acevedo	7	F
262	30/03/2020	Virgen del pilar	54	M
263	30/04/2020	Colonia acevedo	7	F
264	30/04/2020	Virgen del pilar	54	M
265	31/05/2020	Cristo Rey	8	M

Figura Nro 17 Datos epidemiológicos 2020 (continuación)



	Fecha	Ubicación	Edad	Género
266	01/01/2021	Virgen del pilar	33	M
267	01/02/2021	San Jose	67	F
268	01/03/2021	virgen del pilar	33	F
269	01/03/2021	virgen del pilar	67	M
270	01/04/2021	Cerro Costa	20	F
271	01/04/2021	Cerro Costa	22	F
272	01/04/2021	Cerro costa	67	M
273	01/04/2021	virgen del pilar	33	F
274	01/05/2021	San Jose	19	M
275	01/06/2021	Pablo VI	64	F
276	01/06/2021	Pablo VI	70	M
277	01/08/2021	San Geronimo	27	M
278	01/11/2021	San Jose ita moroti	27	M
279	02/03/2021	Pablo VI	14	F
280	02/03/2021	San Geronimo	38	F
281	02/03/2021	San Geronimo	39	M
282	02/04/2021	Pablo VI	14	F
283	02/04/2021	San Geronimo	38	F
284	02/04/2021	San Geronimo	39	M
285	02/06/2021	Pablo VI	50	F
286	02/09/2021	Virgen del pilar	12	F
287	02/09/2021	Virgen del pilar	16	F
288	02/09/2021	Virgen del pilar	33	F
289	02/10/2021	San Miguel	27	F
290	03/03/2021	San Isidro	22	F
291	03/09/2021	Virgen del pilar	16	F
292	04/03/2021	San Geronimo	11	M
293	04/03/2021	san isidro	20	F
294	04/04/2021	San Geronimo	11	M
295	04/10/2021	San Miguel	25	F
296	05/03/2021	San Isidro	87	M

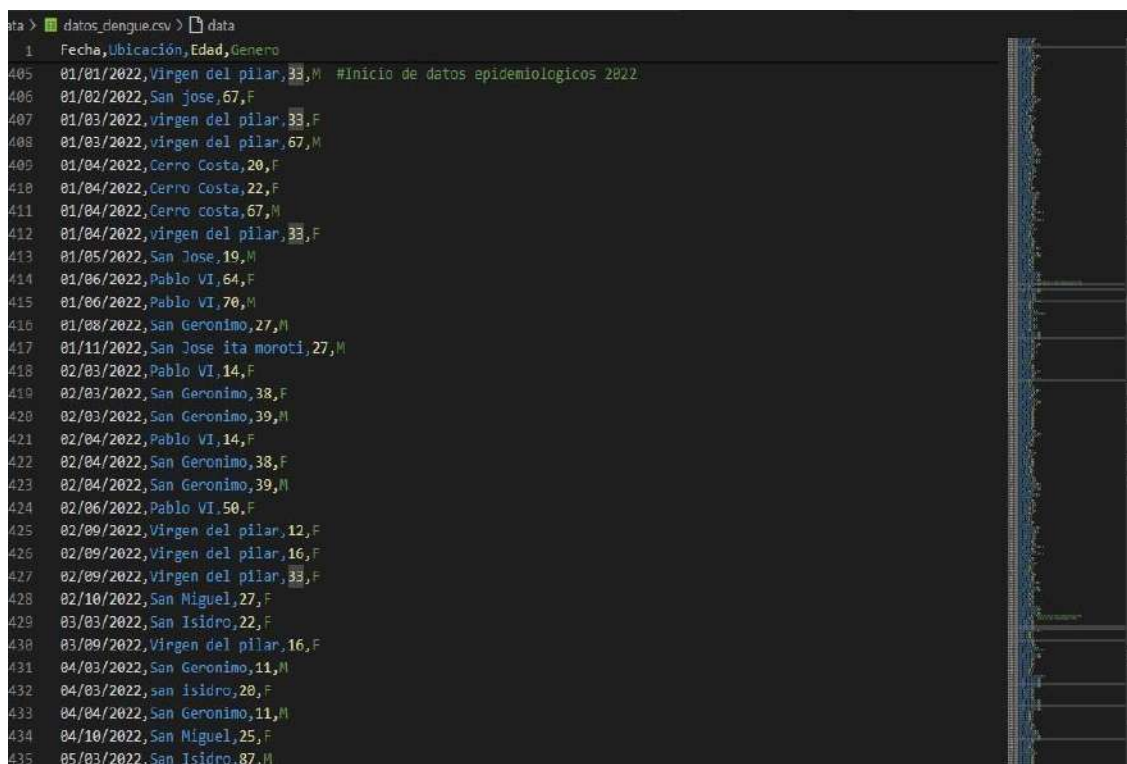
Figura Nro 18 Datos epidemiológicos 2021



```
data > datos_dengue.csv > data
```

	Fecha,Ubicación,Edad,Genero
374	23/03/2021, San Gabriel, 41, M
375	23/03/2021, Virgen del pilar, 71, F
376	23/04/2021, Cristo Rey, 66, F
377	23/04/2021, San Gabriel, 41, M
378	23/04/2021, Virgen del pilar, 71, F
379	23/05/2021, San Jose, 9, M
380	24/09/2021, San Isidro, 35, M
381	25/01/2021, Pablo VI, 23, F
382	25/11/2021, Potrero alto, 41, F
383	27/01/2021, Pablo VI, 22, F
384	27/02/2021, San Jose, 18, F
385	27/05/2021, San Jose, 18, M
386	27/10/2021, San Miguel, 35, F
387	27/12/2021, San Gabriel, 30, F
388	28/03/2021, Cristo Rey, 48, F
389	28/04/2021, Cristo Rey, 48, F
390	28/07/2021, colonia acevedo, 43, F
391	28/10/2021, San Jose, 13, F
392	28/12/2021, colonia acevedo, 43, F
393	29/01/2021, Virgen del pilar, 55, M
394	29/03/2021, Cristo Rey, 44, M
395	29/03/2021, San Jose, 59, M
396	29/04/2021, Cristo Rey, 44, M
397	29/04/2021, San Jose, 59, M
398	29/09/2021, San Isidro, 17, M
399	29/10/2021, San Jose, 23, M
400	30/03/2021, Colonia acevedo, 7, F
401	30/03/2021, Virgen del pilar, 54, M
402	30/04/2021, Colonia acevedo, 7, F
403	30/04/2021, Virgen del pilar, 54, M
404	31/05/2021, Cristo Rey, 8, M #Finalizacion de datos epidemiologicos 2021

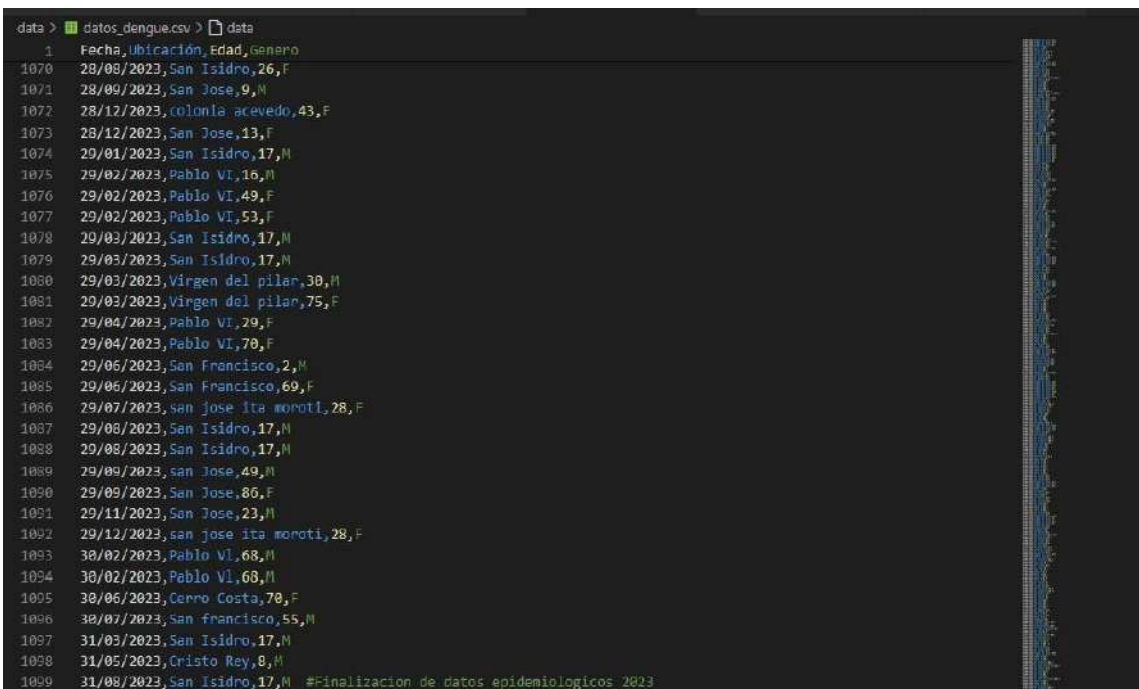
Figura Nro 19 Datos epidemiológicos 2021 (continuación)



```
data > datos_dengue.csv > data
```

	Fecha,Ubicación,Edad,Genero
405	01/01/2022, Virgen del pilar, 33, M #Inicio de datos epidemiologicos 2022
406	01/02/2022, San Jose, 67, F
407	01/03/2022, Virgen del pilar, 33, F
408	01/03/2022, Virgen del pilar, 67, M
409	01/04/2022, Cerro Costa, 20, F
410	01/04/2022, Cerro Costa, 22, F
411	01/04/2022, Cerro costa, 67, M
412	01/04/2022, Virgen del pilar, 33, F
413	01/05/2022, San Jose, 19, M
414	01/06/2022, Pablo VI, 64, F
415	01/06/2022, Pablo VI, 70, M
416	01/08/2022, San Geronimo, 27, M
417	01/11/2022, San Jose ita moroti, 27, M
418	02/03/2022, Pablo VI, 14, F
419	02/03/2022, San Geronimo, 38, F
420	02/03/2022, San Geronimo, 39, M
421	02/04/2022, Pablo VI, 14, F
422	02/04/2022, San Geronimo, 38, F
423	02/04/2022, San Geronimo, 39, M
424	02/06/2022, Pablo VI, 50, F
425	02/09/2022, Virgen del pilar, 12, F
426	02/09/2022, Virgen del pilar, 16, F
427	02/09/2022, Virgen del pilar, 33, F
428	02/10/2022, San Miguel, 27, F
429	03/03/2022, San Isidro, 22, F
430	03/09/2022, Virgen del pilar, 16, F
431	04/03/2022, San Geronimo, 11, M
432	04/03/2022, San Isidro, 20, F
433	04/04/2022, San Geronimo, 11, M
434	04/10/2022, San Miguel, 25, F
435	05/03/2022, San Isidro, 87, M

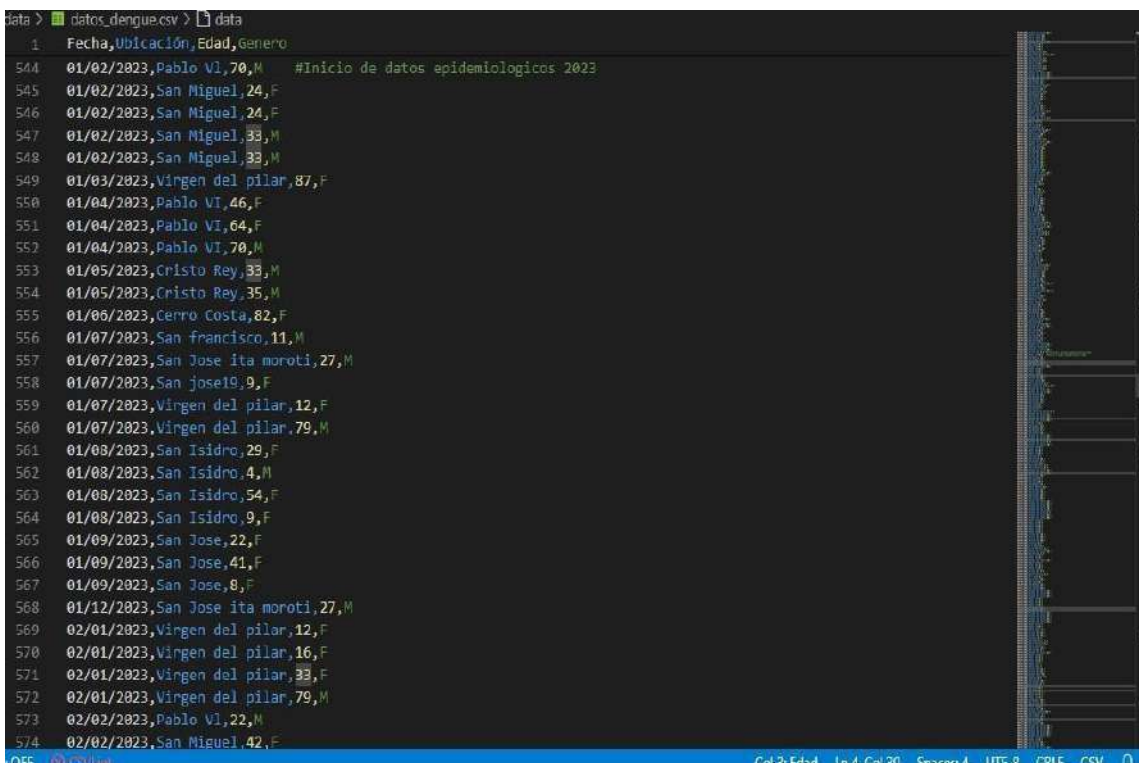
Figura Nro 20 Datos epidemiológicos 2022



	Fecha	Ubicación	Edad	Genero
1070	28/08/2022	San Isidro	26	F
1071	28/09/2022	San Jose	9	M
1072	28/12/2022	colonia acevedo	43	F
1073	28/12/2022	San Jose	13	F
1074	29/01/2023	San Isidro	17	M
1075	29/02/2023	Pablo VI	16	M
1076	29/02/2023	Pablo VI	49	F
1077	29/02/2023	Pablo VI	53	F
1078	29/03/2023	San Isidro	17	M
1079	29/03/2023	San Isidro	17	M
1080	29/03/2023	Virgen del pilar	30	M
1081	29/03/2023	Virgen del pilar	75	F
1082	29/04/2023	Pablo VI	29	F
1083	29/04/2023	Pablo VI	70	F
1084	29/06/2023	San Francisco	2	M
1085	29/06/2023	San Francisco	69	F
1086	29/07/2023	San Jose ita moroti	28	F
1087	29/08/2023	San Isidro	17	M
1088	29/08/2023	San Isidro	17	M
1089	29/09/2023	San Jose	49	M
1090	29/09/2023	San Jose	86	F
1091	29/11/2023	San Jose	23	M
1092	29/12/2023	San Jose ita moroti	28	F
1093	30/02/2023	Pablo VI	68	M
1094	30/02/2023	Pablo VI	68	M
1095	30/06/2023	Cerro Costa	70	F
1096	30/07/2023	San francisco	55	M
1097	31/03/2023	San Isidro	17	M
1098	31/05/2023	Cristo Rey	8	M
1099	31/08/2023	San Isidro	17	M

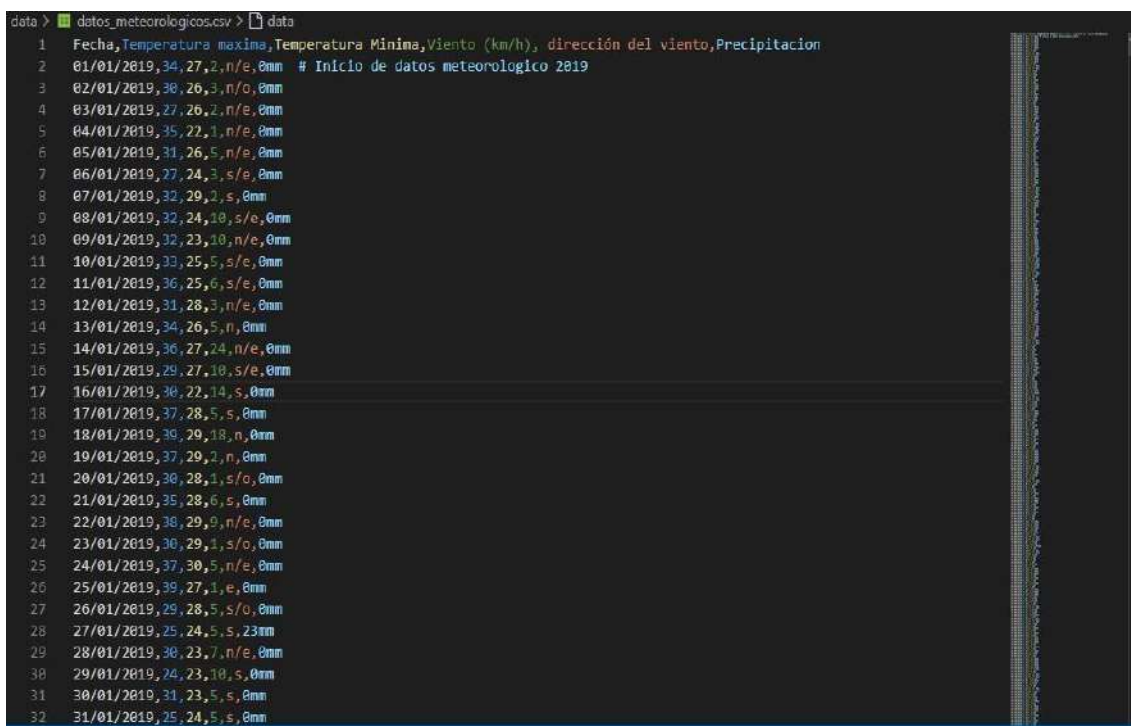
#Finalizacion de datos epidemiologicos 2022

Figura Nro 21 Datos epidemiológicos 2022 (continuación)



	Fecha	Ubicación	Edad	Genero
544	01/02/2023	Pablo VI	70	M
545	01/02/2023	San Miguel	24	F
546	01/02/2023	San Miguel	24	F
547	01/02/2023	San Miguel	33	M
548	01/02/2023	San Miguel	33	M
549	01/03/2023	Virgen del pilar	87	F
550	01/04/2023	Pablo VI	46	F
551	01/04/2023	Pablo VI	64	F
552	01/04/2023	Pablo VI	70	M
553	01/05/2023	Cristo Rey	33	M
554	01/05/2023	Cristo Rey	35	M
555	01/06/2023	Cerro Costa	82	F
556	01/07/2023	San Francisco	11	M
557	01/07/2023	San Jose ita moroti	27	M
558	01/07/2023	San Jose	19	F
559	01/07/2023	Virgen del pilar	12	F
560	01/07/2023	Virgen del pilar	79	M
561	01/08/2023	San Isidro	29	F
562	01/08/2023	San Isidro	4	M
563	01/08/2023	San Isidro	54	F
564	01/08/2023	San Isidro	9	F
565	01/09/2023	San Jose	22	F
566	01/09/2023	San Jose	41	F
567	01/09/2023	San Jose	8	F
568	01/12/2023	San Jose ita moroti	27	M
569	02/01/2023	Virgen del pilar	12	F
570	02/01/2023	Virgen del pilar	16	F
571	02/01/2023	Virgen del pilar	33	F
572	02/01/2023	Virgen del pilar	79	M
573	02/02/2023	Pablo VI	22	M
574	02/02/2023	San Miguel	42	F

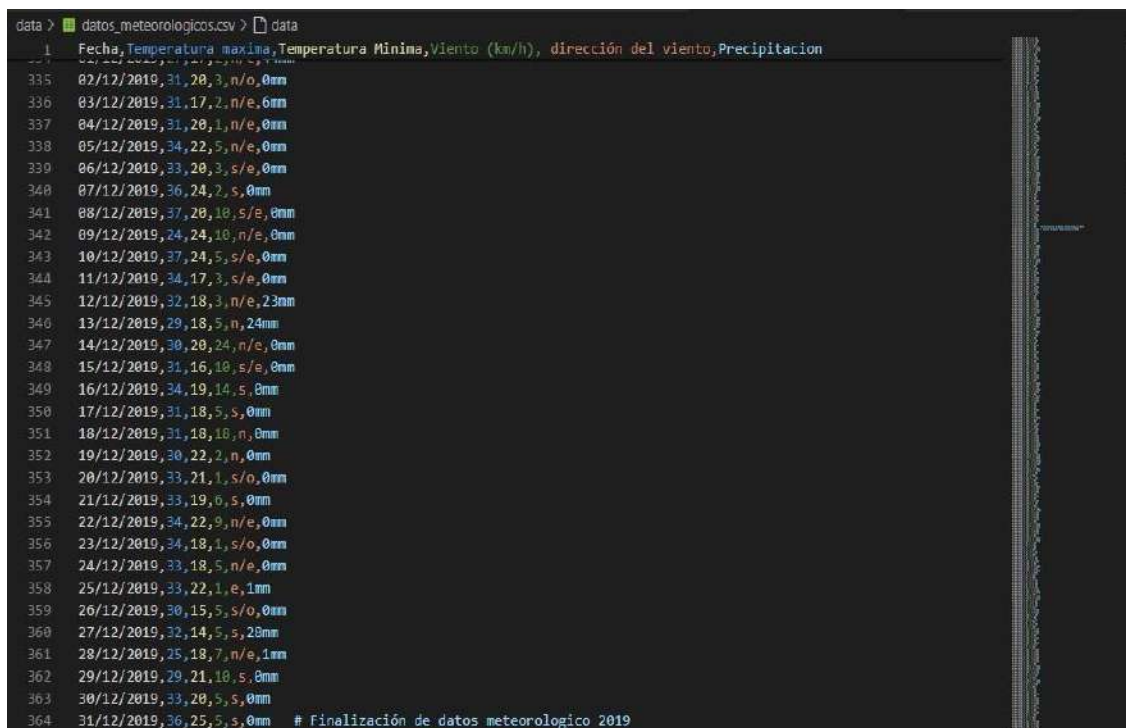
Figura Nro 22 Datos epidemiológicos 2023



```

data > datos_meteorologicos.csv > data
1 Fecha,Temperatura maxima,Temperatura Minima,Viento (km/h), dirección del viento,Precipitacion
2 01/01/2019,34,27,2,n/e,0mm # Inicio de datos meteorologico 2019
3 02/01/2019,36,26,3,n/o,0mm
4 03/01/2019,27,26,2,n/e,0mm
5 04/01/2019,35,22,1,n/e,0mm
6 05/01/2019,31,26,5,n/e,0mm
7 06/01/2019,27,24,3,s/e,0mm
8 07/01/2019,32,29,2,s,0mm
9 08/01/2019,32,24,10,s/e,0mm
10 09/01/2019,32,23,10,n/e,0mm
11 10/01/2019,33,25,5,s/e,0mm
12 11/01/2019,36,25,6,s/e,0mm
13 12/01/2019,31,28,3,n/e,0mm
14 13/01/2019,34,26,5,n,0mm
15 14/01/2019,36,27,24,n/e,0mm
16 15/01/2019,29,27,10,s/e,0mm
17 16/01/2019,36,22,14,s,0mm
18 17/01/2019,37,28,5,s,0mm
19 18/01/2019,36,29,18,n,0mm
20 19/01/2019,37,29,2,n,0mm
21 20/01/2019,36,28,1,s/o,0mm
22 21/01/2019,35,28,6,s,0mm
23 22/01/2019,36,29,9,n/e,0mm
24 23/01/2019,36,29,1,s/o,0mm
25 24/01/2019,37,30,5,n/e,0mm
26 25/01/2019,39,27,1,e,0mm
27 26/01/2019,29,28,5,s/o,0mm
28 27/01/2019,25,24,5,s,23mm
29 28/01/2019,36,23,7,n/e,0mm
30 29/01/2019,24,23,10,s,0mm
31 30/01/2019,31,23,5,s,0mm
32 31/01/2019,25,24,5,s,0mm
    
```

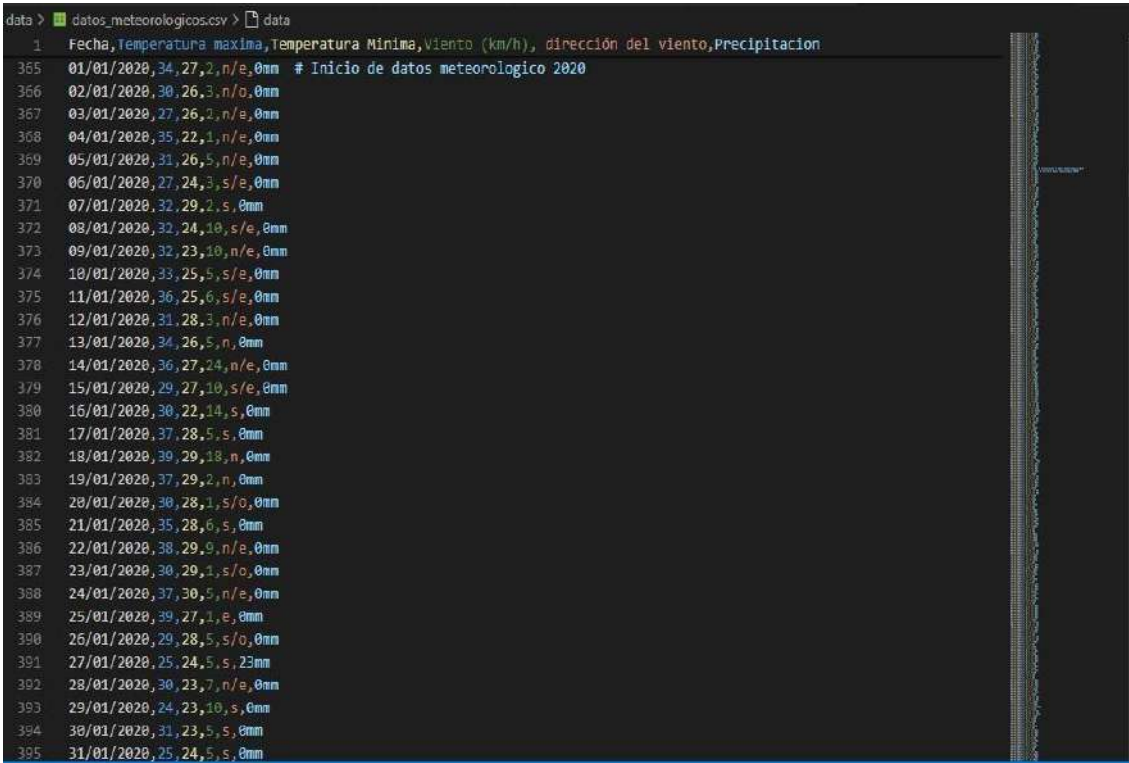
Figura Nro 23 Data base del proyecto meteorológico 2019



```

data > datos_meteorologicos.csv > data
335 02/12/2019,31,28,3,n/o,0mm
336 03/12/2019,31,17,2,n/e,6mm
337 04/12/2019,31,20,1,n/e,0mm
338 05/12/2019,34,22,5,n/e,0mm
339 06/12/2019,33,20,3,s/e,0mm
340 07/12/2019,36,24,2,s,0mm
341 08/12/2019,37,20,10,s/e,0mm
342 09/12/2019,24,24,10,n/e,0mm
343 10/12/2019,37,24,5,s/e,0mm
344 11/12/2019,34,17,3,s/e,0mm
345 12/12/2019,32,18,3,n/e,23mm
346 13/12/2019,29,18,5,n,24mm
347 14/12/2019,30,20,24,n/e,0mm
348 15/12/2019,31,16,10,s/e,0mm
349 16/12/2019,34,19,14,s,0mm
350 17/12/2019,31,18,5,s,0mm
351 18/12/2019,31,18,10,n,0mm
352 19/12/2019,30,22,2,n,0mm
353 20/12/2019,33,21,1,s/o,0mm
354 21/12/2019,33,19,6,s,0mm
355 22/12/2019,34,22,9,n/e,0mm
356 23/12/2019,34,18,1,s/o,0mm
357 24/12/2019,33,18,5,n/e,0mm
358 25/12/2019,33,22,1,e,1mm
359 26/12/2019,30,15,5,s/o,0mm
360 27/12/2019,32,14,5,s,20mm
361 28/12/2019,25,18,7,n/e,1mm
362 29/12/2019,29,21,10,s,0mm
363 30/12/2019,33,20,5,s,0mm
364 31/12/2019,36,25,5,s,0mm # Finalización de datos meteorologico 2019
    
```

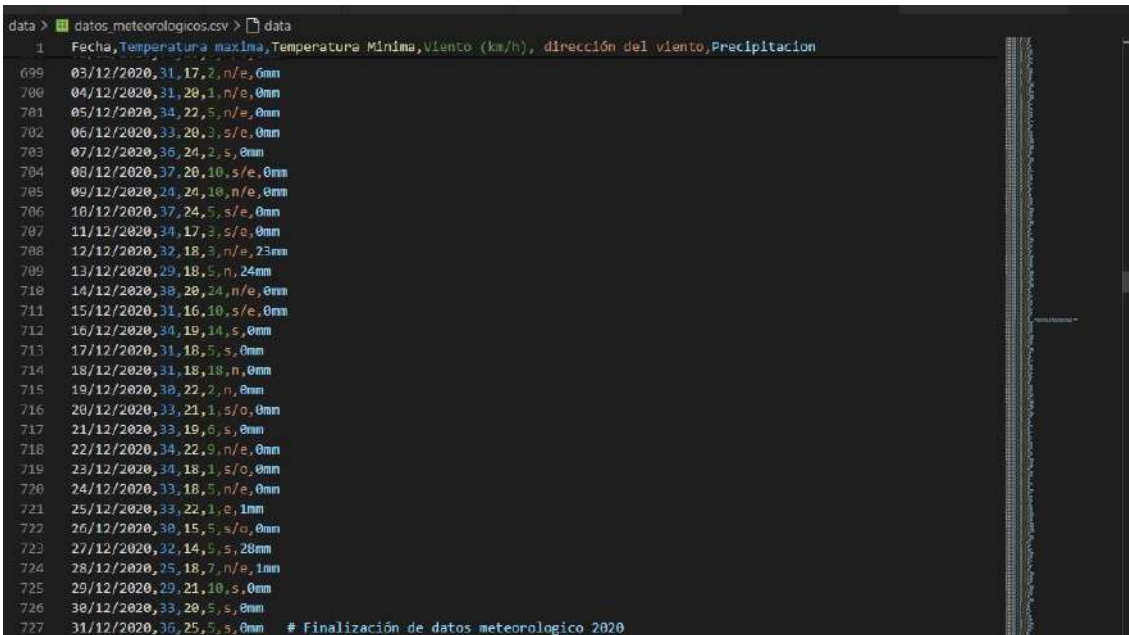
Figura Nro 24 Data base del proyecto meteorológico 2019 (continuación)



```
data > datos_meteorologicos.csv > data
```

	Fecha	Temperatura máxima	Temperatura Mínima	Viento (km/h)	dirección del viento	Precipitación
365	01/01/2020	34,27,2	n/e	0mm		
366	02/01/2020	30,26,3	n/o	0mm		
367	03/01/2020	27,26,2	n/e	0mm		
368	04/01/2020	35,22,1	n/e	0mm		
369	05/01/2020	31,26,5	n/e	0mm		
370	06/01/2020	27,24,3	s/e	0mm		
371	07/01/2020	32,29,2	s	0mm		
372	08/01/2020	32,24,10	s/e	0mm		
373	09/01/2020	32,23,10	n/e	0mm		
374	10/01/2020	33,25,5	s/e	0mm		
375	11/01/2020	36,25,6	s/e	0mm		
376	12/01/2020	31,28,3	n/e	0mm		
377	13/01/2020	34,26,5	n	0mm		
378	14/01/2020	36,27,24	n/e	0mm		
379	15/01/2020	29,27,10	s/e	0mm		
380	16/01/2020	30,22,14	s	0mm		
381	17/01/2020	37,28,5	s	0mm		
382	18/01/2020	39,29,18	n	0mm		
383	19/01/2020	37,29,2	n	0mm		
384	20/01/2020	30,28,1	s/o	0mm		
385	21/01/2020	35,28,6	s	0mm		
386	22/01/2020	38,29,9	n/e	0mm		
387	23/01/2020	30,29,1	s/o	0mm		
388	24/01/2020	37,30,5	n/e	0mm		
389	25/01/2020	39,27,1	e	0mm		
390	26/01/2020	29,28,5	s/o	0mm		
391	27/01/2020	25,24,5	s	23mm		
392	28/01/2020	30,23,7	n/e	0mm		
393	29/01/2020	24,23,10	s	6mm		
394	30/01/2020	31,23,5	s	0mm		
395	31/01/2020	25,24,5	s	0mm		

Figura Nro 25 Data base del proyecto meteorológico 2020



```
data > datos_meteorologicos.csv > data
```

	Fecha	Temperatura máxima	Temperatura Mínima	Viento (km/h)	dirección del viento	Precipitación
699	03/12/2020	31,17,2	n/e	0mm		
700	04/12/2020	31,20,1	n/e	0mm		
701	05/12/2020	34,22,5	n/e	0mm		
702	06/12/2020	33,20,3	s/e	0mm		
703	07/12/2020	36,24,2	s	0mm		
704	08/12/2020	37,20,10	s/e	0mm		
705	09/12/2020	24,24,10	n/e	0mm		
706	10/12/2020	37,24,5	s/e	0mm		
707	11/12/2020	34,17,3	s/e	0mm		
708	12/12/2020	32,18,3	n/e	23mm		
709	13/12/2020	29,18,5	n	24mm		
710	14/12/2020	30,20,24	n/e	0mm		
711	15/12/2020	31,16,10	s/e	0mm		
712	16/12/2020	34,19,14	s	0mm		
713	17/12/2020	31,18,5	s	0mm		
714	18/12/2020	31,18,18	n	0mm		
715	19/12/2020	30,22,2	n	0mm		
716	20/12/2020	33,21,1	s/o	0mm		
717	21/12/2020	33,19,6	s	0mm		
718	22/12/2020	34,22,9	n/e	0mm		
719	23/12/2020	34,18,1	s/o	0mm		
720	24/12/2020	33,18,5	n/e	0mm		
721	25/12/2020	33,22,1	e	1mm		
722	26/12/2020	30,15,5	s/o	0mm		
723	27/12/2020	32,14,5	s	28mm		
724	28/12/2020	25,18,7	n/e	1mm		
725	29/12/2020	29,21,10	s	0mm		
726	30/12/2020	33,20,5	s	0mm		
727	31/12/2020	36,25,5	s	0mm		

Figura Nro 26 Data base del proyecto meteorológico 2020 (continuación)

```

a > datos_meteorologicos.csv > data
1  Fecha,Temperatura maxima,Temperatura Minima,Viento (km/h), dirección del viento,Precipitacion
28 01/01/2021,34,27,2,n/e,0mm # Inicio de datos meteorológico 2021
29 02/01/2021,30,26,3,n/o,0mm
30 03/01/2021,27,26,2,n/e,0mm
31 04/01/2021,35,22,1,n/e,0mm
32 05/01/2021,31,26,5,n/e,0mm
33 06/01/2021,27,24,3,s/e,0mm
34 07/01/2021,32,29,2,s,0mm
35 08/01/2021,32,24,10,s/e,0mm
36 09/01/2021,32,23,10,n/e,0mm
37 10/01/2021,33,25,5,s/e,0mm
38 11/01/2021,36,25,0,s/e,0mm
39 12/01/2021,31,28,3,n/e,0mm
40 13/01/2021,34,26,5,n,0mm
41 14/01/2021,36,27,24,n/e,0mm
42 15/01/2021,29,27,10,s/e,0mm
43 16/01/2021,30,22,14,s,0mm
44 17/01/2021,37,28,5,s,0mm
45 18/01/2021,39,29,18,n,0mm
46 19/01/2021,37,29,2,n,0mm
47 20/01/2021,30,28,1,s/o,0mm
48 21/01/2021,35,28,6,s,0mm
49 22/01/2021,38,29,9,n/e,0mm
50 23/01/2021,30,29,1,s/o,0mm
51 24/01/2021,37,30,5,n/e,0mm
52 25/01/2021,39,27,1,e,0mm
53 26/01/2021,29,28,5,s/o,0mm
54 27/01/2021,25,24,5,s,23mm
55 28/01/2021,30,23,7,n/e,0mm
56 29/01/2021,24,23,10,s,0mm
57 30/01/2021,31,23,5,s,0mm
58 31/01/2021,25,24,5,s,0mm

```

Figura Nro 27 Data base del proyecto meteorológico 2021

```

data > datos_meteorologicos.csv > data
1  Fecha,Temperatura maxima,Temperatura Minima,Viento (km/h), dirección del viento,Precipitacion
1060 01/12/2021,27,17,2,n/e,44mm
1061 02/12/2021,31,20,3,n/o,0mm
1062 03/12/2021,31,17,2,n/e,6mm
1063 04/12/2021,31,20,1,n/e,0mm
1064 05/12/2021,34,22,5,n/e,0mm
1065 06/12/2021,33,20,3,s/e,0mm
1066 07/12/2021,36,24,2,s,0mm
1067 08/12/2021,37,20,10,s/e,0mm
1068 09/12/2021,24,24,10,n/e,0mm
1069 10/12/2021,37,24,5,s/e,0mm
1070 11/12/2021,34,17,3,s/e,0mm
1071 12/12/2021,32,18,3,n/e,23mm
1072 13/12/2021,29,18,5,n,24mm
1073 14/12/2021,30,20,24,n/e,0mm
1074 15/12/2021,31,16,10,s/e,0mm
1075 16/12/2021,34,19,14,s,0mm
1076 17/12/2021,31,18,5,s,0mm
1077 18/12/2021,31,18,18,n,0mm
1078 19/12/2021,30,22,2,n,0mm
1079 20/12/2021,33,21,1,s/o,0mm
1080 21/12/2021,33,19,6,s,0mm
1081 22/12/2021,24,22,9,n/e,0mm
1082 23/12/2021,34,18,1,s/o,0mm
1083 24/12/2021,33,18,5,n/e,0mm
1084 25/12/2021,33,22,1,e,1mm
1085 26/12/2021,30,15,5,s/o,0mm
1086 27/12/2021,32,14,5,s,28mm
1087 28/12/2021,25,18,7,n/e,1mm
1088 29/12/2021,29,21,10,s,0mm
1089 30/12/2021,33,20,5,s,0mm
1090 31/12/2021,36,25,5,s,0mm # Finalización de datos meteorológico 2021

```

Figura Nro 28 Data base del proyecto meteorológico 2021 (continuación)

```
data > datos_meteorologicos.csv > data
```

	1	Fecha	Temperatura maxima	Temperatura Minima	Viento (km/h)	direccion del viento	Precipitacion
1091	01/01/2022	34	27	2	n/e	0mm	
1092	02/01/2022	30	26	3	n/o	0mm	
1093	03/01/2022	27	26	2	n/e	0mm	
1094	04/01/2022	35	22	1	n/e	0mm	
1095	05/01/2022	31	26	5	n/e	0mm	
1096	06/01/2022	27	24	3	s/e	0mm	
1097	07/01/2022	32	29	2	s	0mm	
1098	08/01/2022	32	24	10	s/e	0mm	
1099	09/01/2022	32	23	10	n/e	0mm	
1100	10/01/2022	33	25	5	s/e	0mm	
1101	11/01/2022	36	25	6	s/e	0mm	
1102	12/01/2022	31	28	3	n/e	0mm	
1103	13/01/2022	34	26	5	n	0mm	
1104	14/01/2022	36	27	24	n/e	0mm	
1105	15/01/2022	29	27	10	s/e	0mm	
1106	16/01/2022	30	22	14	s	0mm	
1107	17/01/2022	37	28	5	s	0mm	
1108	18/01/2022	39	29	18	n	0mm	
1109	19/01/2022	37	29	2	n	0mm	
1110	20/01/2022	30	28	1	s/o	0mm	
1111	21/01/2022	35	28	6	s	0mm	
1112	22/01/2022	38	29	9	n/e	0mm	
1113	23/01/2022	30	29	1	s/o	0mm	
1114	24/01/2022	37	30	5	n/e	0mm	
1115	25/01/2022	39	27	1	e	0mm	
1116	26/01/2022	29	28	5	s/o	0mm	
1117	27/01/2022	25	24	5	s	13mm	
1118	28/01/2022	30	23	7	n/e	0mm	
1119	29/01/2022	24	23	10	s	0mm	
1120	30/01/2022	31	23	5	s	0mm	
1121	31/01/2022	25	24	5	s	0mm	

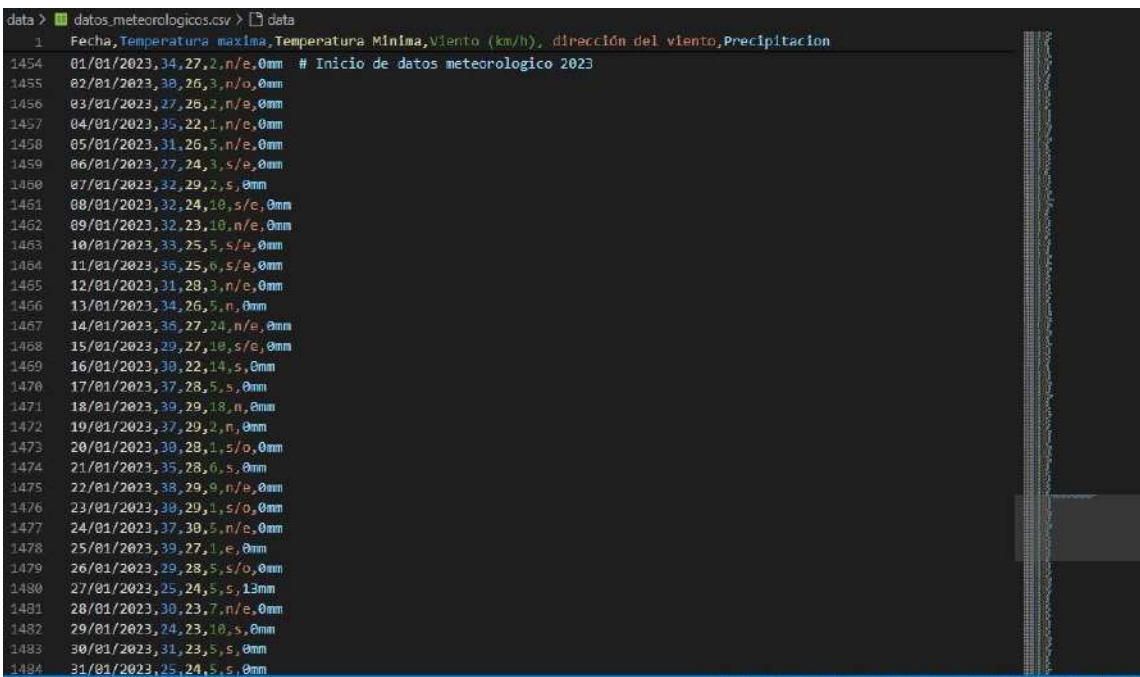
Figura Nro 29 Data base del proyecto meteorológico 2022

```
data > datos_meteorologicos.csv > data
```

	1	Fecha	Temperatura maxima	Temperatura Minima	Viento (km/h)	direccion del viento	Precipitacion
1443	01/12/2022	27	17	2	n/e	0mm	
1424	02/12/2022	31	20	3	n/o	0mm	
1425	03/12/2022	31	17	2	n/e	6mm	
1426	04/12/2022	31	20	1	n/e	0mm	
1427	05/12/2022	34	22	5	n/e	0mm	
1428	06/12/2022	33	20	3	s/e	0mm	
1429	07/12/2022	36	24	2	s	0mm	
1430	08/12/2022	37	20	10	s/e	0mm	
1431	09/12/2022	24	24	10	n/e	0mm	
1432	10/12/2022	37	24	5	s/e	0mm	
1433	11/12/2022	34	17	3	s/e	0mm	
1434	12/12/2022	32	18	3	n/e	23mm	
1435	13/12/2022	29	18	5	n	24mm	
1436	14/12/2022	30	20	24	n/e	0mm	
1437	15/12/2022	31	16	10	s/e	0mm	
1438	16/12/2022	34	19	14	s	0mm	
1439	17/12/2022	31	18	5	s	0mm	
1440	18/12/2022	31	18	18	n	0mm	
1441	19/12/2022	30	22	2	n	0mm	
1442	20/12/2022	33	21	1	s/o	0mm	
1443	21/12/2022	33	19	6	s	0mm	
1444	22/12/2022	34	22	9	n/e	0mm	
1445	23/12/2022	34	18	1	s/o	0mm	
1446	24/12/2022	33	18	5	n/e	0mm	
1447	25/12/2022	33	22	1	e	1mm	
1448	26/12/2022	30	15	5	s/o	0mm	
1449	27/12/2022	32	14	5	s	28mm	
1450	28/12/2022	25	18	7	n/e	1mm	
1451	29/12/2022	29	21	10	s	0mm	
1452	30/12/2022	33	20	5	s	0mm	
1453	31/12/2022	36	25	5	s	0mm	

Finalización de datos meteorologico 2022

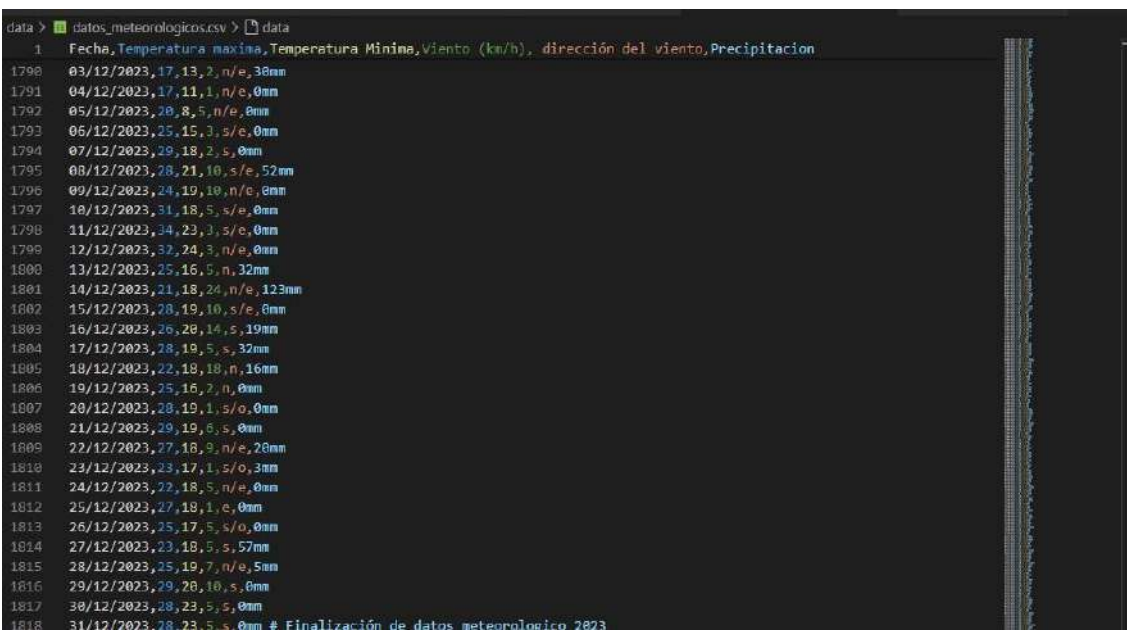
Figura Nro 30 Data base del proyecto meteorológico 2022 (continuación)



```
data > datos_meteorologicos.csv > data
```

	1	Fecha	Temperatura máxima	Temperatura Mínima	Viento (km/h)	dirección del viento	Precipitación
1454	01/01/2023	34	27	2	n/e	0mm	# Inicio de datos meteorológico 2023
1455	02/01/2023	30	26	3	n/o	0mm	
1456	03/01/2023	27	26	2	n/e	0mm	
1457	04/01/2023	35	22	1	n/e	0mm	
1458	05/01/2023	31	26	5	n/e	0mm	
1459	06/01/2023	27	24	3	s/e	0mm	
1460	07/01/2023	32	29	2	s	0mm	
1461	08/01/2023	32	24	10	s/e	0mm	
1462	09/01/2023	32	23	10	n/e	0mm	
1463	10/01/2023	33	25	5	s/e	0mm	
1464	11/01/2023	30	25	0	s/e	0mm	
1465	12/01/2023	31	28	3	n/e	0mm	
1466	13/01/2023	34	26	5	n	0mm	
1467	14/01/2023	36	27	24	n/e	0mm	
1468	15/01/2023	29	27	10	s/e	0mm	
1469	16/01/2023	30	22	14	s	0mm	
1470	17/01/2023	37	28	5	s	0mm	
1471	18/01/2023	30	29	18	n	0mm	
1472	19/01/2023	37	29	2	n	0mm	
1473	20/01/2023	30	28	1	s/o	0mm	
1474	21/01/2023	35	28	6	s	0mm	
1475	22/01/2023	38	29	9	n/e	0mm	
1476	23/01/2023	30	29	1	s/o	0mm	
1477	24/01/2023	37	30	5	n/e	0mm	
1478	25/01/2023	39	27	1	e	0mm	
1479	26/01/2023	29	28	5	s/o	0mm	
1480	27/01/2023	25	24	5	s	13mm	
1481	28/01/2023	30	23	7	n/e	0mm	
1482	29/01/2023	24	23	10	s	0mm	
1483	30/01/2023	31	23	5	s	0mm	
1484	31/01/2023	25	24	5	s	0mm	

Figura Nro 31 Data base del proyecto meteorológico 2023



```
data > datos_meteorologicos.csv > data
```

	1	Fecha	Temperatura máxima	Temperatura Mínima	Viento (km/h)	dirección del viento	Precipitación
1790	03/12/2023	17	13	2	n/e	30mm	
1791	04/12/2023	17	11	1	n/e	0mm	
1792	05/12/2023	20	8	5	n/e	0mm	
1793	06/12/2023	25	15	3	s/e	0mm	
1794	07/12/2023	29	18	2	s	0mm	
1795	08/12/2023	28	21	10	s/e	52mm	
1796	09/12/2023	24	19	10	n/e	0mm	
1797	10/12/2023	31	18	5	s/e	0mm	
1798	11/12/2023	34	23	3	s/e	0mm	
1799	12/12/2023	32	24	3	n/e	0mm	
1800	13/12/2023	25	16	5	n	32mm	
1801	14/12/2023	21	18	24	n/e	123mm	
1802	15/12/2023	20	19	10	s/e	0mm	
1803	16/12/2023	26	20	14	s	19mm	
1804	17/12/2023	28	19	5	s	32mm	
1805	18/12/2023	22	18	18	n	16mm	
1806	19/12/2023	25	16	2	n	0mm	
1807	20/12/2023	20	19	1	s/o	0mm	
1808	21/12/2023	20	19	6	s	0mm	
1809	22/12/2023	27	18	9	n/e	20mm	
1810	23/12/2023	23	17	1	s/o	3mm	
1811	24/12/2023	22	18	5	n/e	0mm	
1812	25/12/2023	27	18	1	e	0mm	
1813	26/12/2023	25	17	5	s/o	0mm	
1814	27/12/2023	23	18	5	s	57mm	
1815	28/12/2023	25	19	7	n/e	5mm	
1816	29/12/2023	29	20	10	s	0mm	
1817	30/12/2023	28	23	5	s	0mm	
1818	31/12/2023	28	23	5	s	0mm	# Finalización de datos meteorológico 2023

Figura Nro 32 Data base del proyecto meteorológico 2023 (continuación)

Los dos conjuntos de datos principales utilizados en esta simulación fueron **datos_dengue.csv** y **datos_meteorologico.csv**. El primero, **datos_dengue.csv**, fue construido utilizando información epidemiológica recopilada durante los años 2022, 2023 y 2024, incluyendo información sobre la fecha de inicio de los síntomas, los lugares exactos dentro de la ciudad de Santa Rosa Misiones y algunos de los datos demográficos de los pacientes, como su edad y su género, entre otros. Los datos mencionados son vitales para identificar la propagación del dengue y para la predicción de un modelo sólido y confiable.

El archivo **datos_meteorologico.csv** contiene información acerca de las condiciones meteorológicas asociadas a la propagación del dengue, con columnas de fecha, temperatura máxima y mínima, velocidad del viento en km/h, dirección del viento y precipitaciones. La inclusión de estas variables meteorológicas en la presente base ha permitido documentar correlaciones apreciables entre las variables climáticas y la presencia de la especie *Aedes aegypti*, principal vector del dengue.

Para realizar la simulación, se ha utilizado Python y técnicas de ML para extender los datos históricos epidemiológicos y meteorológicos y, por consiguiente, prever la propagación futura del dengue. Desde el comienzo del ciclo, se cargan los datos reales que luego se exponen dentro de un gráfico que registra la trayectoria desde el inicio hasta la fecha de los casos reportados y de las condiciones climáticas de los últimos años, otorgando una representación integral de las tendencias históricas.

La visualización inicial de los datos otorga la base de la simulación dinámica, la cual se realiza utilizando técnicas de ML que predicen si el dengue se expandiría basándose en la historia y los cambios esperados en las condiciones climáticas. A medida que cambian las condiciones climáticas y se ejecuta la simulación, la visualización también cambia en tiempo real, continuamente actualizada para ofrecer una presentación visual dinámica de la propagación esperada de dengue en Santa Rosa Misiones.

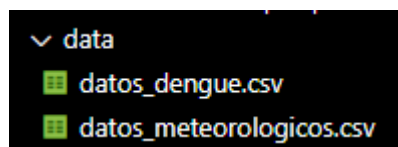


Figura Nro 33 *Data base epidemiológico y meteorológico*

342. Etapa 2: Diseño e Integración

En esta segunda etapa se llevan a cabo las actividades siguientes:

3421. Definición de las herramientas específicas

En esta fase, se escogieron y definieron las herramientas y tecnologías clave para la simulación de la propagación del dengue, ya evaluadas en la fase exploratoria, lo cual incluyó:

- **Lenguaje de programación:** Se usó Python por su flexibilidad y extensa variedad de bibliotecas disponibles para ML y procesamiento de datos.
- **Fuentes de datos:** El uso de datos epidemiológicos otorgados por el hospital, junto con datos meteorológicos históricos, sirvió para alimentar y optimizar el modelo predictivo.
- **Algoritmos de predicción:** Se seleccionaron redes neuronales LSTM (Long Short-Term Memory), un modelo de aprendizaje profundo especialmente adecuado para el análisis de secuencias temporales, debido a su capacidad para realizar predicciones precisas sobre la propagación del dengue.
- **Herramientas de visualización:** El empleo de Matplotlib fue para la visualización en tiempo real de los casos simulados de dengue y tkinter, de forma que se desarrolla una interfaz gráfica que facilita el control y monitoreo de la simulación.

Herramientas utilizadas:

```
simulation.py
import tkinter as tk
from tkinter import ttk, messagebox, filedialog
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
import matplotlib.pyplot as plt
import pandas as pd
import numpy as np
import threading
import time
from sklearn.preprocessing import MinMaxScaler
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, LSTM, Dropout
from tensorflow.keras.callbacks import EarlyStopping
import matplotlib.dates as mdates
from sklearn.metrics import mean_squared_error, mean_absolute_error
import locale
import sys
import time
```

Figura Nro 34 Importaciones

- **Import tkinter as tk y from tkinter import ttk, messagebox, filedialog:** Tkinter es una biblioteca estándar de Python que permite la creación de interfaces gráficas de usuario (GUI, por sus siglas en inglés). **Ttk** (Themed Tkinter), por otro lado, es un módulo de Tkinter que ayuda a crear widgets con estilos actualizados y personalizables, incluyendo botones y etiquetas. Se utilizó el módulo que contiene el widget **messagebox** y **filedialog** para mostrar cuadros de diálogo, incorporando ventanas emergentes para mensajes y de selección de archivos, respectivamente, para la mejor experiencia del usuario.
- **From matplotlib.backends.backend_tkagg import FigureCanvasTkAgg:** Matplotlib es una biblioteca utilizada para generar gráficos en Python, y el módulo **FigureCanvasTkAgg** permite integrar estos gráficos directamente en una ventana de Tkinter, facilitando la visualización de datos en aplicaciones con interfaz gráfica.
- **Import matplotlib.pyplot as plt:** Matplotlib es la biblioteca más ampliamente utilizada para la creación de gráficos en Python. El submódulo **pyplot**, importado comúnmente como **plt**, proporciona una interfaz simplificada para generar diversos tipos de gráficos, como líneas, barras e histogramas, facilitando así la visualización de datos.
- **Import pandas as pd:** Pandas es una biblioteca esencial en Python para la manipulación y análisis de datos. El alias común **pd** se utiliza para trabajar con estructuras de datos como **DataFrames**, que representan tablas de datos bidimensionales y permiten operaciones eficientes de filtrado, agregación y transformación de datos.
- **Import numpy as np:** NumPy es una biblioteca fundamental para realizar cálculos numéricos y manipulación de matrices y vectores en Python. Se utiliza comúnmente el alias **np** para facilitar el trabajo con sus funciones y operaciones en contextos de procesamiento de datos y computación científica.
- **Import threading:** En Python, el uso del módulo **threading** permite a los usuarios realizar varias tareas concurrentemente, lo que puede ser de gran utilidad para cualquier aplicación que requiera procesamiento paralelo.

- **Import time: Time** es una biblioteca estándar de Python para realizar funciones relacionadas con el tiempo, como pausas temporales, obtener la hora actual y realizar muchas otras operaciones de control de tiempo en el script de código.
- **From sklearn.preprocessing import MinMaxScaler:** Scikit-learn (**sklearn**) es una biblioteca de ML en Python. **MinMaxScaler** es una herramienta dentro de esta biblioteca que permite escalar (normalizar) los datos para que se encuentren dentro de un rango especificado, usualmente entre 0 y 1, facilitando así el procesamiento uniforme en algoritmos de Aprendizaje Automático.
- **From tensorflow.keras.models import Sequential:** **TensorFlow** constituye una biblioteca de código abierto diseñada para la creación y entrenamiento de modelos de ML. Keras, por su parte, incluida dentro de TensorFlow, es una API de alto nivel que simplifica la construcción de redes neuronales. Mientras que, el modelo **Sequential** en Keras permite definir una arquitectura de red neuronal en la que las capas se agregan de manera secuencial, proporcionando una estructura lineal y clara para desarrollar modelos.
- **From tensorflow.keras.layers import Dense, LSTM, Dropout:**
 - **Dense:** Una capa completamente conectada en una red neuronal, en la cual cada neurona recibe entradas de todas las neuronas de la capa anterior, permitiendo una combinación integral de las características.
 - **LSTM:** Un tipo de capa de red neuronal recurrente diseñada para manejar secuencias de datos y realizar predicciones a largo plazo, ideal para series temporales y procesamiento de datos secuenciales.
 - **Dropout:** Una técnica utilizada para prevenir el sobreajuste (overfitting), que desactiva aleatoriamente ciertas neuronas durante el entrenamiento, promoviendo así una mayor generalización del modelo.
- **From tensorflow.keras.callbacks import EarlyStopping:** **EarlyStopping** es un callback en Keras que interrumpe el entrenamiento del modelo cuando el rendimiento deja de mejorar, ayudando a prevenir el sobreajuste y optimizando la eficiencia del proceso de entrenamiento.

- **Import matplotlib.dates as mdates: mdates, mdates**, módulo de Matplotlib especializado en la manipulación y formato de fechas en gráficos, facilitando la representación precisa de datos temporales.
- **From sklearn.metrics import mean_squared_error, mean_absolute_error: mean_squared_error (MSE) y mean_absolute_error (MAE)** son métricas de evaluación utilizadas para medir la precisión de un modelo de predicción, proporcionando una referencia sobre el nivel de error en las predicciones del modelo.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Figura Nro 35 Error cuadrático medio (MSE)

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n}$$

Figura Nro 36 Error Absoluto Medio (MAE)

3422. Fases del desarrollo del sistema

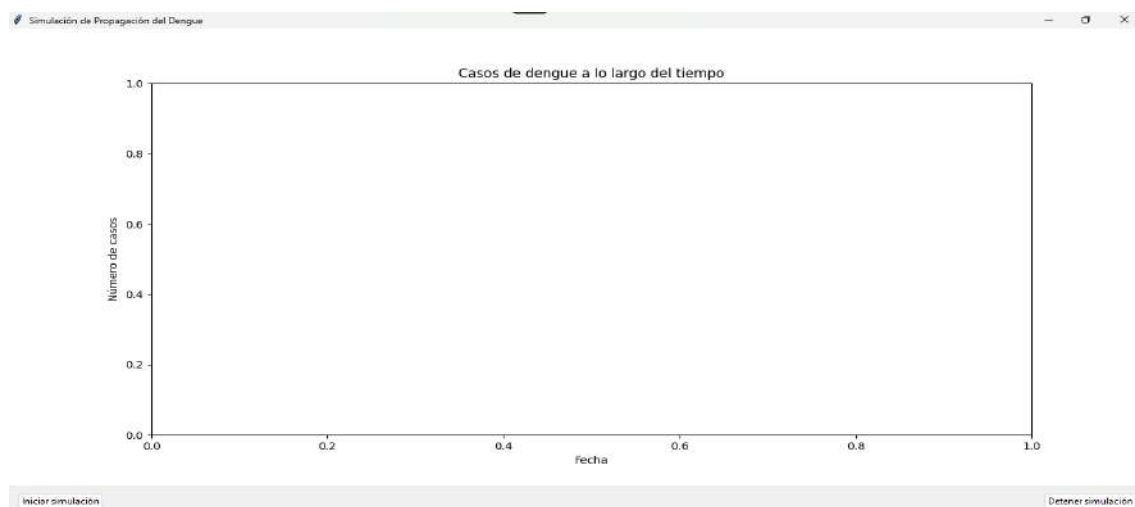


Figura Nro 37 Interfaz gráfica utilizando tkinter

Tkinter se utilizó para crear los elementos de la interfaz como los botones de control para iniciar la simulación. Además, permitió mostrar gráficos interactivos que visualizaban en tiempo real el comportamiento de la propagación del dengue.

- **Iniciar la simulación:**

Activa el modelo de ML para predecir futuros casos de dengue, utilizando datos históricos epidemiológicos y meteorológicos como base.

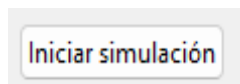


Figura Nro 38 Botón Iniciar Simulación

- **Visualizar gráficos en tiempo real:**

Se producen gráficos que ilustran los casos históricos completados y las expectativas futuras procesadas por el programa. Los datos son acumulados y presentados dinámicamente en Matplotlib, lo que permite visualizar la disminución y elevación del número de casos de dengue.

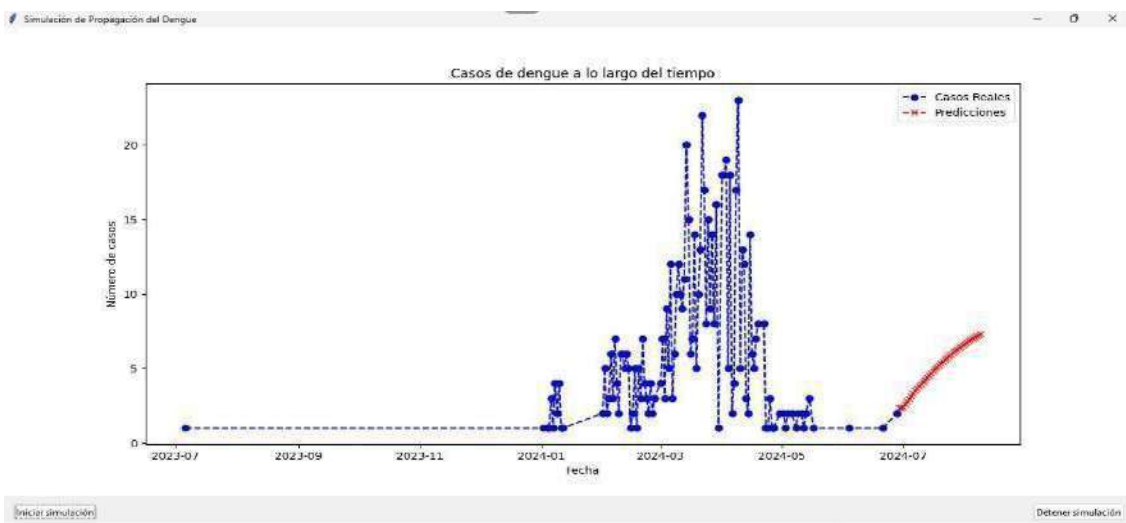


Figura Nro 39 Gráfico de datos reales y predicción

Se observó que el sistema leyó únicamente los datos correspondientes a 2024, en lugar de los datos históricos de 2019 a 2023, por dicha razón, se procedió a ajustar la programación, de manera a asegurar la inclusión y procesamiento correcto de los datos históricos mencionados.

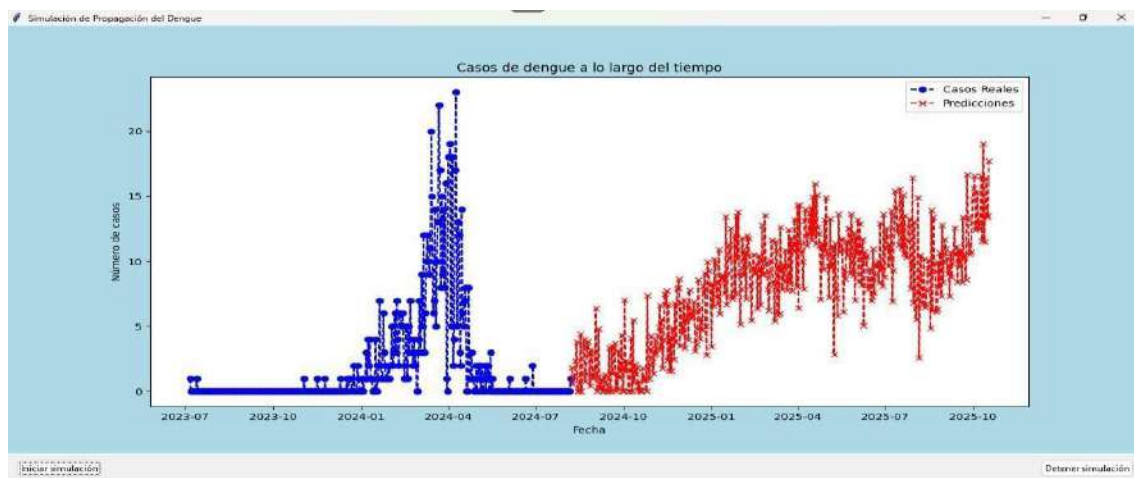


Figura Nro 40 Gráfico haciendo predicción 2023 2024

Asimismo, se añadió un fondo de color celeste para mejorar la visualización temporal, configurando el sistema para procesar correctamente los datos históricos de 2023 y 2024, lo cual permitió representar gráficamente las predicciones actuales. Sin embargo, para optimizar las predicciones de futuros casos de propagación de dengue, aún es necesario incorporar los datos de 2019, ya que estos permitirán que el modelo de ML ajuste sus proyecciones con mayor precisión, mejorando así el tiempo de detección de futuros problemas.

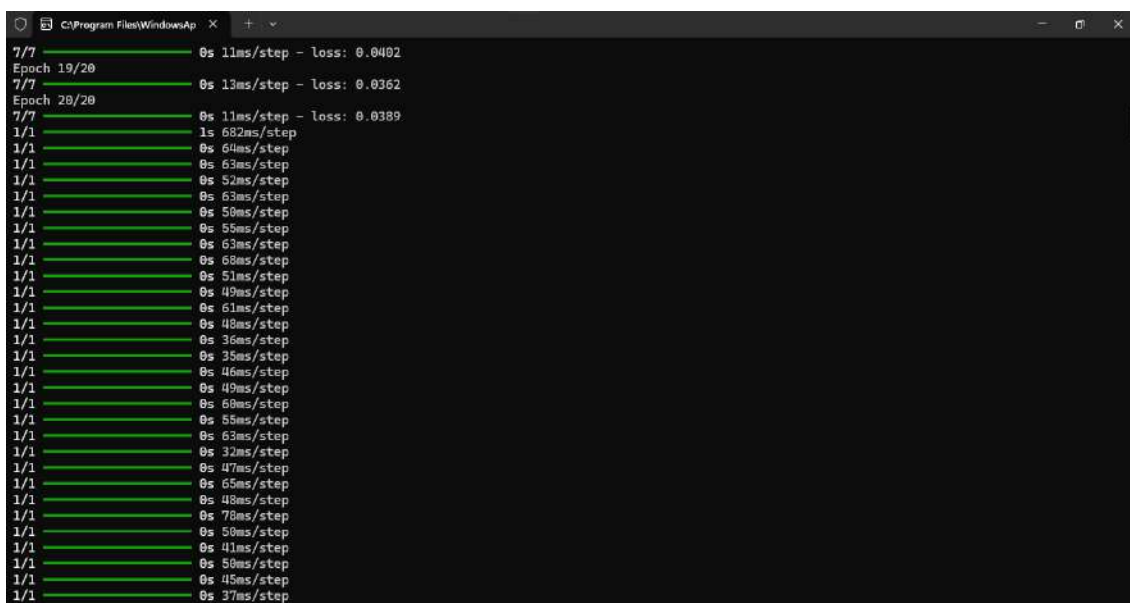


Figura Nro 41 Símbolo de sistema sin error

Los mensajes no indicaron errores críticos; únicamente presentaron advertencias relacionadas con optimizaciones y buenas prácticas. El modelo parece estar entrenando correctamente, según se observó en la métrica de pérdida.

El prototipo se ha construido para otorgar una interfaz intuitiva y fácil de usar para las autoridades sanitarias y el público en general interesado en cómo se propaga el dengue. El diseño de la interfaz de usuario facilita la navegación y la interpretación de los resultados, al mismo tiempo, el gráfico dinámico ayuda a prever el futuro a partir de los rangos temporales pasados de la enfermedad. Un gráfico estable podría no sugerir nuevos casos, en cambio, ante un posible brote, se reflejarán aumentos en las proyecciones de casos.

- **Integración del modelo de predicción con la interfaz gráfica:**

El modelo de redes neuronales LSTM se integró en la interfaz gráfica para garantizar que los datos procesados hayan sido reflejados en tiempo real, entrenándose con datos históricos de casos de dengue y condiciones meteorológicas, como la temperatura y precipitación, con vistas a predecir posibles brotes de la enfermedad en el futuro.

Las predicciones generadas por el modelo se muestran en la interfaz, y los gráficos responden dinámicamente a las proyecciones del modelo. Este proceso de integración fue fundamental para crear una simulación coherente que combine datos reales con predicciones futuras.

- **Corrección de fechas**

El siguiente punto crítico es el manejo de los datos epidemiológicos y meteorológicos, en los cuales es necesario tener un estricto control para que todas las fechas estén correctamente formateadas y ordenadas, de manera a garantizar la precisión de las predicciones y simulaciones. Cualquier error en el registro de fechas puede provocar inconsistencias en el análisis temporal, afectando la fiabilidad de los resultados dados por el modelo de ML. Dado lo anterior, se implementó un proceso de corrección y validación de fechas.

```

corregir_fecha.py
1 import pandas as pd
2 from datetime import datetime
3
4 # Función para corregir las fechas
5 def corregir_fecha(fecha):
6     try:
7         # Intentar convertir la fecha al formato correcto
8         return pd.to_datetime(fecha, format='%d/%m/%Y', errors='coerce').strftime('%d/%m/%Y')
9     except:
10         return None
11
12 # Cargar mis datos en un DataFrame
13 df = pd.read_csv('datos_dengue.csv')
14
15 # Aplicar la corrección a la columna de fechas
16 df['Fecha'] = df['Fecha'].apply(corregir_fecha)
17
18 # Guardar los datos corregidos en un nuevo archivo CSV
19 df.to_csv('tus_datos_corregidos.csv', index=False)
20
21 print("Las fechas han sido corregidas y el nuevo archivo se ha guardado como 'tus_datos_corregidos.csv'.")

```

Figura Nro 42 Corrección de fecha

- **Leída de datos históricos 2019 a 2023**

Se logró que la IA lea y procese los datos epidemiológicos y meteorológicos correspondientes al año 2019, a partir de este avance, el siguiente paso consistió en reajustar el código y optimizar el modelo para habilitar la predicción de casos de dengue en 2024. Una vez generadas estas predicciones, se compararon con los datos reales de 2024 para evaluar la precisión y eficacia del modelo de ML, mejorar las futuras predicciones y validar la confiabilidad de la simulación en tiempo real.

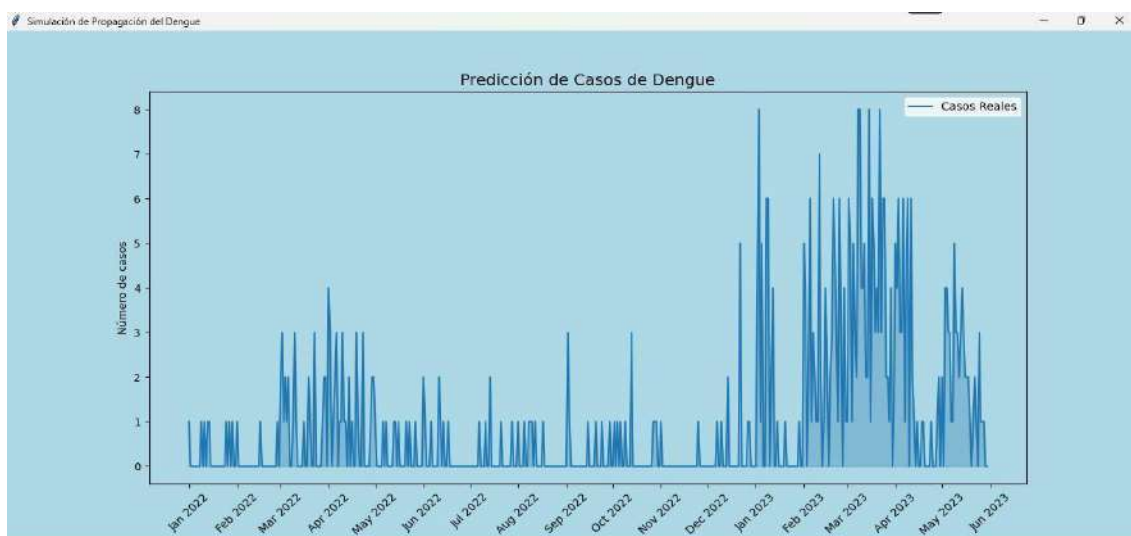


Figura Nro 43 Prueba de leído de datos reales

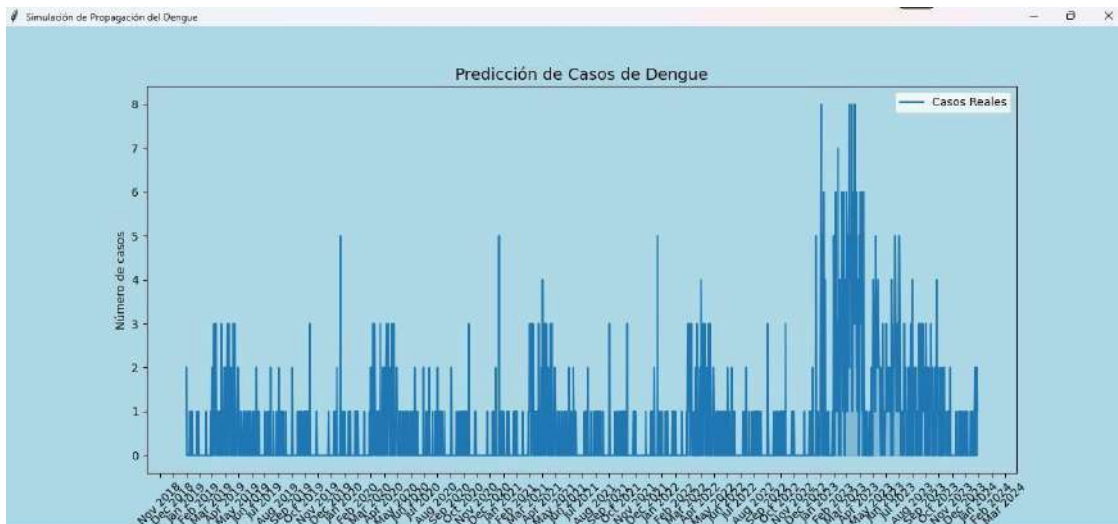


Figura Nro 44 *Prueba de leído de datos reales 2019 a 2023 para predicción del 2024 con problema de fechas justas*

En este punto, se observó que las fechas estaban muy próximas entre sí, lo que dificultó una visualización clara de los datos, por lo tanto, se realizó un ajuste enfocado en un entrenamiento avanzado de ML que permita visualizar las predicciones de manera anual, de esta manera, al seleccionar un año específico, como 2024, el sistema muestra exclusivamente las predicciones de enero a diciembre de dicho año. Las proyecciones se generarán utilizando los datos históricos epidemiológicos y meteorológicos como principales factores predictivos.

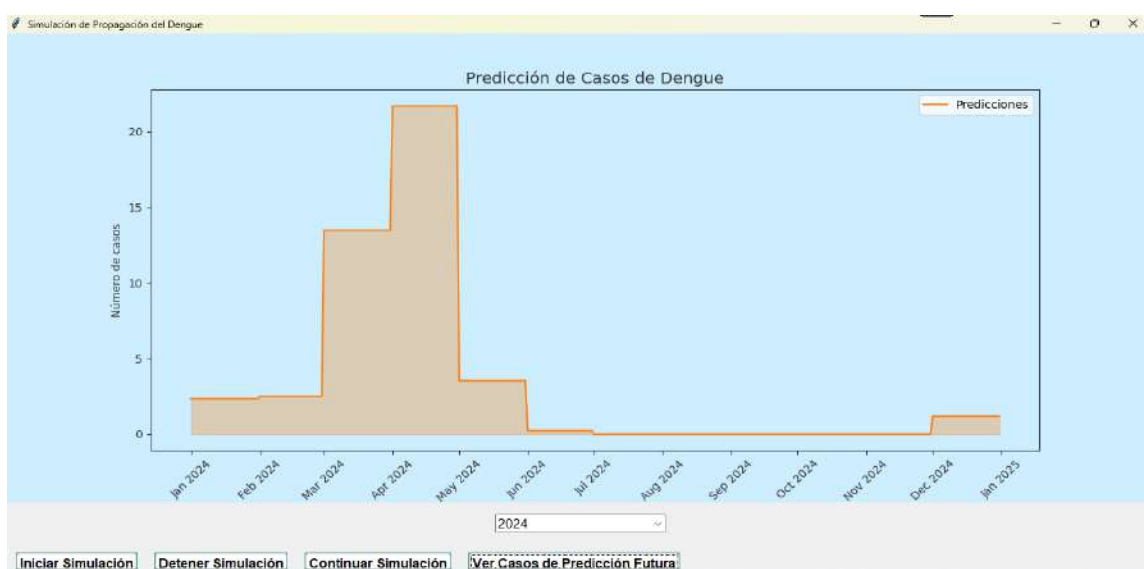


Figura Nro 45 *Prueba de ajuste*

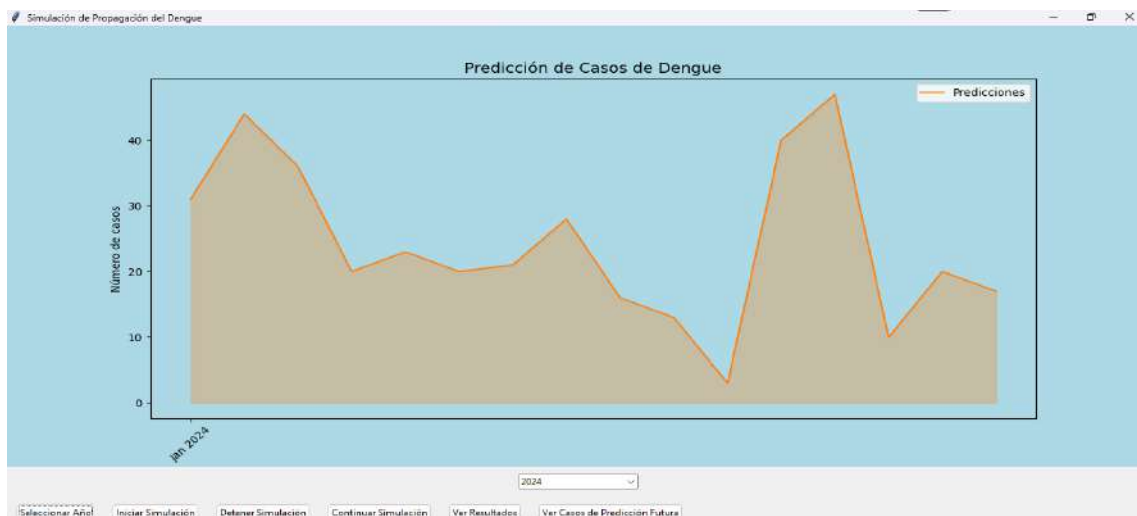


Figura Nro 46 Visualización de entrenamiento de predicción para el año 2024

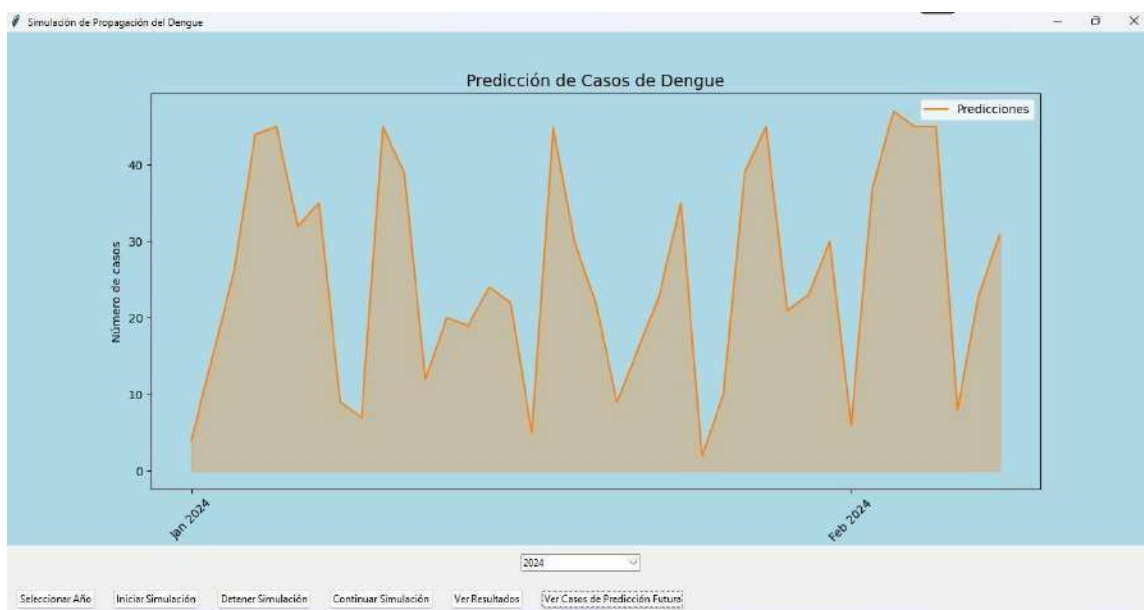


Figura Nro 47 Prueba de predicción para año 2024

Se incorporó posteriormente un fondo más profesional y se añadieron botones específicos que optimizaron el funcionamiento y la usabilidad del sistema, permitiendo al usuario interactuar de manera más intuitiva y eficiente.

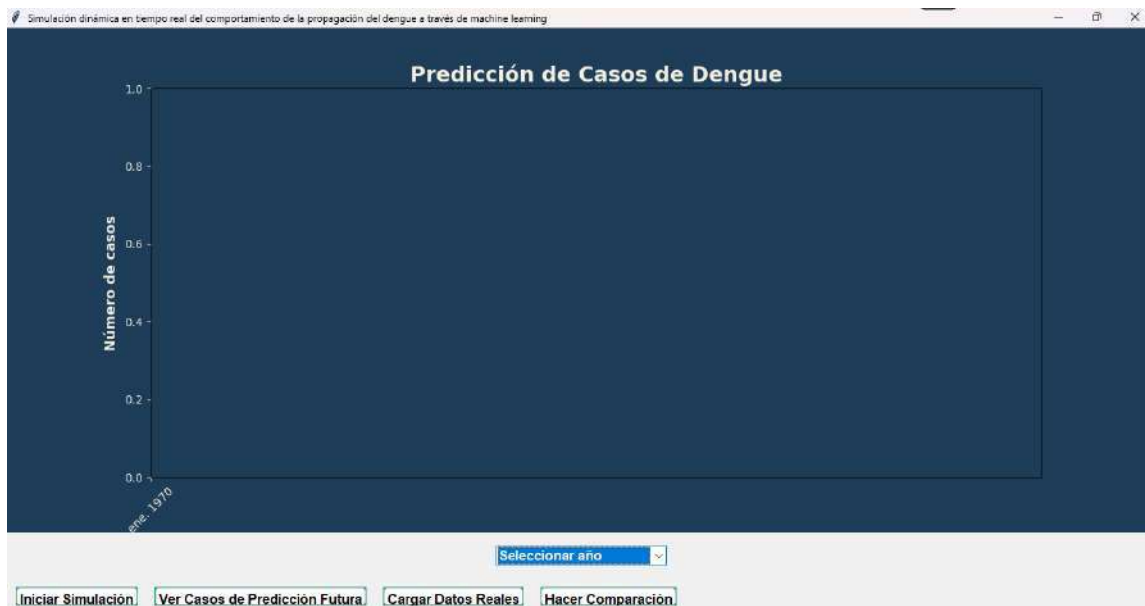


Figura Nro 48 Visualización profesional de la simulación

3.4.2.3. Implementación de los botones en la simulación

En este apartado, se implementó un conjunto de botones diseñado para permitir de manera eficiente y amigable las siguientes acciones: selección de años futuros, inicio de la simulación, visualización de casos proyectados, carga de datos reales desde un archivo Excel y comparación entre datos predichos y reales. El enfoque mencionado optimiza la experiencia del usuario, facilitando la navegación y el uso intuitivo del sistema.

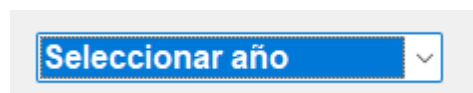


Figura Nro 49 Botón selección de año

La funcionalidad de selección de año implementada en la interfaz de usuario permite al sistema simular y predecir casos de dengue en función del año especificado por el usuario, organizando la lógica de predicción en torno a los datos históricos cargados y las proyecciones futuras, optimizando la experiencia del usuario y facilitando un uso intuitivo del simulador.

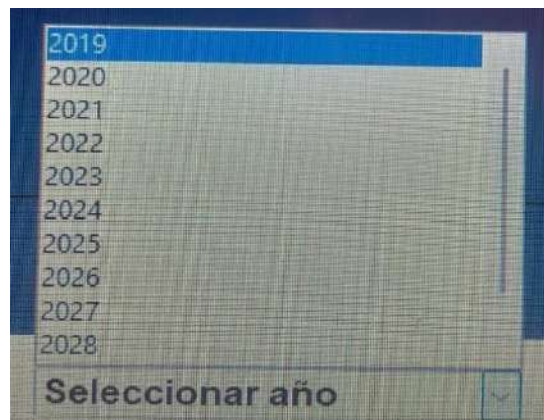


Figura Nro 50 Selección de años futuros

Menú desplegable de selección de año

El sistema utiliza un menú desplegable **ttk.Combobox** que permite al usuario seleccionar un año específico para ejecutar la simulación, el control está configurado para listar años desde 2019 hasta 2030, abarcando tanto aquellos con datos históricos como años futuros en los que se generarán predicciones basadas en el modelo de ML implementado.

Cabe destacar que, el menú desplegable está configurado en modo de solo lectura (*state="readonly"*), lo que impide al usuario ingresar texto directamente y le obliga a seleccionar un año de la lista predefinida, de manera a garantizar que el sistema procese únicamente entradas válidas, reduciendo así el riesgo de errores de ejecución derivados de entradas no reconocidas.

Validación de la selección de año

Antes de iniciar la simulación, el sistema valida que el usuario haya seleccionado un año correctamente. En este sentido, si se intenta proceder sin elegir un año, se muestra un mensaje de error indicando la necesidad de realizar esta selección para continuar. La verificación es fundamental para evitar la ejecución de simulaciones incompletas o basadas en datos inexistentes, garantizando un flujo de programa robusto y libre de errores al simular datos para años no seleccionados. El mensaje de error informa al usuario del paso pendiente, asegurando una interacción clara y precisa.

Diferenciación entre años históricos y futuros

Años históricos (2019-2023): Para estos años, al ser seleccionados por el usuario, el sistema busca y grafica los datos reales disponibles, permitiendo así analizar el comportamiento histórico de la propagación del dengue en función de registros previamente documentados.

Años futuros (2024 en adelante): Al seleccionar un año futuro, el sistema aplica un enfoque predictivo, generando proyecciones de casos mensuales mediante un modelo de ML entrenado con los datos históricos disponibles, tanto epidemiológicos como meteorológicos.

Botón: "Iniciar Simulación"

Este botón inicia el proceso de simulación para el año seleccionado en el menú desplegable, así, al activarse, genera una predicción mensual de los casos de dengue en el año indicado, utilizando los datos históricos y el modelo de ML basado en la red neuronal LSTM. Los resultados de la simulación se presentan en un gráfico que muestra la evolución de los casos a lo largo del tiempo.

Detalles:

- En caso de que el usuario no seleccione un año, se muestra un mensaje de error solicitando que realice la selección, para garantizar que la simulación se ejecute únicamente cuando el año ha sido correctamente seleccionado.

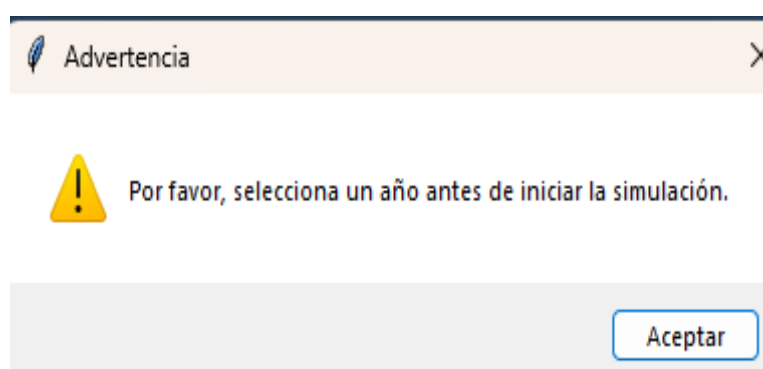


Figura Nro 51 *Mensaje de advertencia*

- Si el año seleccionado es 2024 o posterior, el sistema solicita confirmación al usuario para reiniciar la simulación, lo cual resulta útil cuando se han generado predicciones previas y el usuario desea ejecutar una nueva simulación.

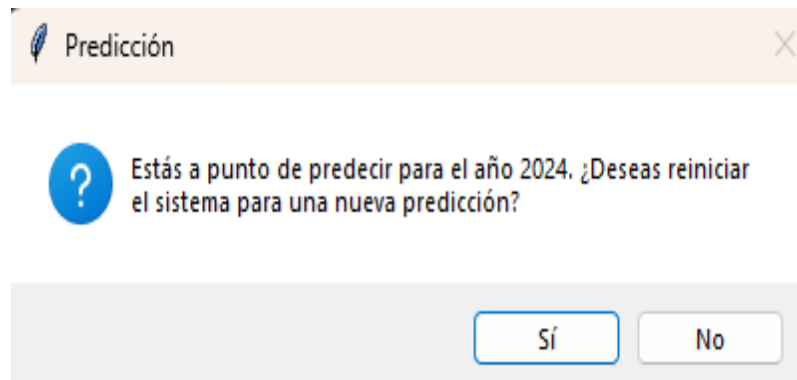


Figura Nro 52 Reinicio de sistema

- En caso de que no existan datos históricos para el año seleccionado, el sistema preguntará al usuario si desea generar una predicción. Si el usuario acepta, se procederá a iniciar el proceso de predicción.

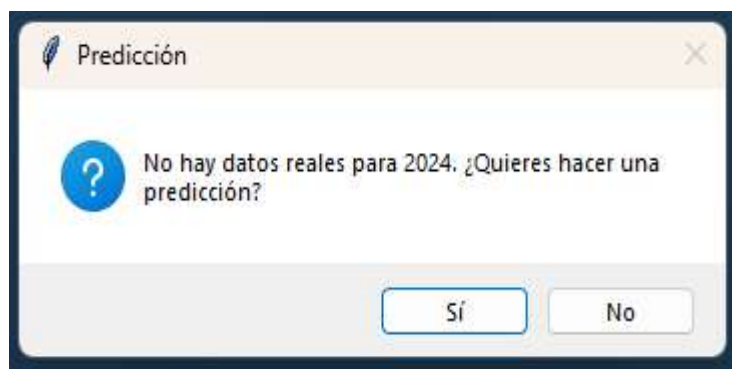


Figura Nro 53 Confirmación para realizar predicción

- Durante la simulación, el sistema genera predicciones mensuales para el año seleccionado y las presenta en una visualización gráfica.

Botón: "Ver Casos de Predicción Futura "

Este botón permite al usuario visualizar las predicciones de casos de dengue en años futuros, al seleccionarlo, se abre una nueva ventana en la que el usuario puede elegir un año específico, y el sistema muestra los casos proyectados para cada mes de ese año.

Tabla Nro 1

Comparación de casos reales con la predicción

Mes	Casos Reales	Predicción
Enero	73	68
Febrero	73	69
Marzo	418	412
Abril	651	646
Mayo	110	105
Junio	7	5
Julio	2	1
Agosto	0	0
Septiembre	0	0
Octubre	0	0
Noviembre	?	0
Diciembre	?	37

Los resultados demuestran que el sistema es capaz de predecir los casos de dengue con un alto grado de exactitud, ajustándose de manera precisa a los patrones observados en los datos históricos. Meses como enero, febrero, marzo y abril presentan variaciones mínimas entre los valores reales y las predicciones, lo que refuerza la confiabilidad del modelo para anticipar la propagación del dengue. Cabe destacar que las mayores variaciones se observan en los meses con menor incidencia de casos.

El sistema, además de generar predicciones con un alto grado de precisión, se adapta eficazmente a las tendencias estacionales y epidemiológicas, lo que lo convierte en una herramienta robusta para la toma de decisiones en salud pública. La integración de técnicas avanzadas de ML, junto con datos históricos y meteorológicos, ha sido fundamental para alcanzar este nivel de exactitud.

Se procede a cargar los datos reales en formato Excel dentro del sistema mediante la funcionalidad "Cargar Datos Reales," la cual permite seleccionar el archivo de Excel correspondiente que contiene los casos reales de dengue recopilados en el Hospital de Santa Rosa, Misiones.

3.5.3. Red Neuronal LSTM (Long Short-Term Memory)

Constituye un tipo de Red Neuronal Recurrente (RNN) diseñada para el procesamiento y modelado de secuencias de datos a lo largo del tiempo. Las arquitecturas LSTM están optimizadas para retener patrones y relaciones en secuencias extensas, superando así las limitaciones de las RNN tradicionales, que enfrentan dificultades debido al problema del desvanecimiento del gradiente.

Características principales de las redes LSTM:

- **Celdas de memoria:** Las LSTM tienen celdas de memoria que pueden almacenar información relevante durante largos períodos de tiempo, lo cual les permite mantener el contexto de los datos secuenciales.
- **Puertas de control:** Las LSTM utilizan tres tipos de "puertas" que controlan el flujo de información:
 - **Puerta de entrada:** Decide qué parte de la nueva información se debe guardar en la celda de memoria.
 - **Puerta de olvido:** Decide qué información en la celda de memoria debe ser olvidada o descartada.
 - **Puerta de salida:** Controla qué parte de la información almacenada en la celda de memoria se debe pasar a la salida en cada paso temporal.
- **Manejo de secuencias:** Las LSTM son particularmente útiles en tareas donde es importante considerar el orden y la dependencia de los datos a lo largo del tiempo. Lo mencionado las hace muy efectivas para aplicaciones como la predicción de series temporales, procesamiento de lenguaje natural (traducción de texto, análisis de sentimientos), y la predicción en secuencias biológicas, entre otras.

3.5.4. Pruebas

Eficiencia del sistema de simulación dinámica del dengue

En el contexto de la simulación dinámica del comportamiento de la propagación del dengue mediante técnicas de ML, la eficiencia del sistema se define por la optimización de los recursos computacionales empleados, incluyendo el tiempo de procesamiento y el uso de memoria. La optimización es esencial para garantizar predicciones precisas y rápidas en tiempo real, ya que el sistema debe gestionar grandes volúmenes de datos epidemiológicos y meteorológicos para generar predicciones de manera continua.

El objetivo principal consiste en minimizar el uso de recursos mientras se asegura la precisión de las predicciones, optimizando así el rendimiento del sistema. La eficiencia del sistema puede ser analizada en términos de la optimización de un proceso repetitivo, donde se persigue mejorar tanto la velocidad de la simulación como la capacidad para gestionar datos históricos y futuros, sin comprometer la exactitud de las predicciones.

Para evaluar el desempeño del sistema, se llevaron a cabo pruebas que consideraron dos métricas fundamentales: el tiempo promedio de respuesta y el consumo de memoria. En este contexto, el tiempo de respuesta se define como el intervalo transcurrido entre la solicitud de una predicción y la obtención de los resultados de la simulación. Para cuantificar este tiempo, se ejecutó el sistema en múltiples ocasiones (aproximadamente 20) para cada conjunto de datos históricos y futuros.

El consumo de memoria, por su parte, es un factor crítico en la simulación, ya que un uso excesivo puede resultar en ralentizaciones significativas en el procesamiento y, en situaciones más severas, incluso provocar fallos del sistema. Un manejo eficiente de la memoria asegura que el sistema opere de manera continua y sin interrupciones, lo que permite al usuario llevar a cabo simulaciones y comparaciones de forma fluida.

Lectura de Datos CSV en el código

La carga de datos se efectúa a partir de un archivo CSV, en el cual las fechas son procesadas y agrupadas por día, el procedimiento permite la visualización en tiempo real del número de casos reportados por fecha, lo que facilita un análisis continuo de la propagación de la enfermedad.

Manejo de errores en la lectura de datos

Durante la lectura del archivo CSV, especificado como `csv_file=r'C:\Users\holad\OneDrive\Escritorio\dengue_simulation_project\data\datos_dengue.csv'`, se identificó un problema relacionado con el manejo de rutas en Python, particularmente cuando estas contienen espacios o caracteres especiales, este inconveniente surge de la interacción entre el sistema operativo y el módulo Tkinter.

En Python, las barras invertidas simples (`\`) se emplean para designar caracteres especiales, lo que puede generar errores al interpretar rutas de archivos. Para corregir esta situación, se optó por utilizar barras invertidas dobles (`\\`), lo que asegura una interpretación correcta de la ruta, el uso del módulo `os.path` de Python permite validar la ruta y garantizar su correcta lectura por el sistema.

```
csv_dengue_file_path = 'C:\\Users\\holad\\OneDrive\\Escritorio\\dengue_simulation_project\\data\\datos_dengue.csv'
csv_meteorological_file_path = 'C:\\Users\\holad\\OneDrive\\Escritorio\\dengue_simulation_project\\data\\datos_meteorologicos.csv'
```

Figura Nro 79 Doble barra invertida para la lectura de la base de datos

El problema se resolvió implementando el uso de barras invertidas dobles (`\\`) en la ruta de acceso al archivo CSV, asegurando así la correcta lectura de los datos.

Error en las fechas en la base de datos

Durante la ejecución del programa, se produjo un error relacionado con el formato de las fechas en la base de datos, el cual arrojó el siguiente mensaje en la consola: `File "strptime.pyx", line 583, in pandas._libs.tslibs.strptime._parse_with_format ValueError: time data "08/003/24" doesn't match format "%d/%m/%Y", at position 94.`

El error señalado indica que los datos de las fechas no coinciden con el formato esperado

```
corregir_fechas.py
1 import pandas as pd
2 from datetime import datetime
3
4 # Función para corregir las fechas
5 def corregir_fecha(fecha):
6     try:
7         # Intentar convertir la fecha al formato correcto
8         return pd.to_datetime(fecha, format='%d/%m/%Y', errors='coerce').strftime('%d/%m/%Y')
9     except:
10         return None
11
12 # Cargar mis datos en un DataFrame
13 df = pd.read_csv('datos_dengue.csv')
14
15 # Aplicar la corrección a la columna de fechas
16 df['Fecha'] = df['Fecha'].apply(corregir_fecha)
17
18 # Guardar los datos corregidos en un nuevo archivo CSV
19 df.to_csv('tus_datos_corregidos.csv', index=False)
20
21 print("Las fechas han sido corregidas y el nuevo archivo se ha guardado como 'tus_datos_corregidos.csv'.")
```

("%d/%m/%Y"), lo que impacta negativamente en el procesamiento del sistema. Para abordar este inconveniente, se implementó una función que corrige automáticamente las fechas mal formateadas en el archivo CSV. La función verifica y ajusta el formato de las fechas antes de proceder con el análisis de los datos, garantizando así que el sistema pueda procesarlas correctamente y sin interrupciones.

Figura Nro 80 *Función de corrección de fechas de forma automática*

3.5.1.1. Primera prueba del sistema

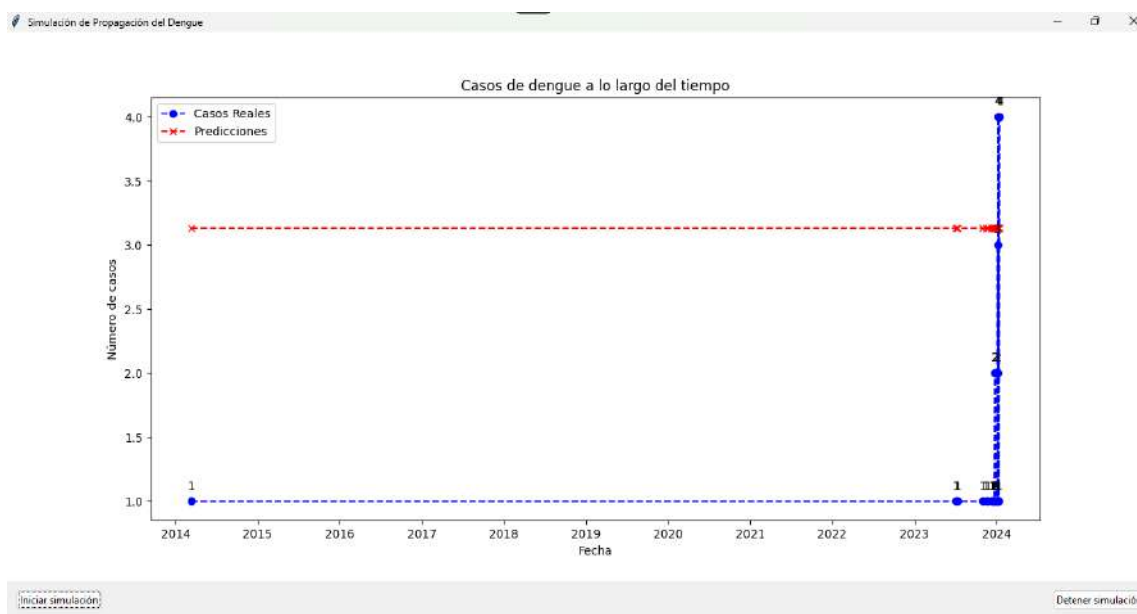


Figura Nro 81 *Primera prueba del sistema*

- **EarlyStopping:** Técnica de optimización que finaliza el entrenamiento de un modelo cuando su rendimiento deja de mejorar, evitando el sobreajuste y reduciendo el tiempo de procesamiento.
- **MinMaxScaler:** Método de preprocesamiento de datos que escala los valores de entrada para que se encuentren en un rango determinado, generalmente entre 0 y 1, lo cual ayuda a mejorar la convergencia y precisión del modelo de ML.
- **API:** Application Programming Interface (Interfaz de Programación de Aplicaciones), conjunto de definiciones y protocolos que permiten la comunicación entre diferentes aplicaciones de software, facilitando el acceso y uso de funciones de un sistema por parte de otros programas.
- **GUI:** Graphical User Interface (Interfaz Gráfica de Usuario), interfaz visual que permite la interacción con el sistema mediante elementos gráficos como ventanas, botones y menús, facilitando el uso del software.
- **Tkinter:** Biblioteca estándar de Python para la creación de GUI. Tkinter proporciona widgets y herramientas que permiten construir interfaces visuales de manera sencilla.
- **ttk (Themed Tkinter):** extensión de Tkinter que ofrece widgets de aspecto moderno y personalizable, mejorando la estética y funcionalidad de las interfaces gráficas en Python.
- **DataFrame:** estructura de datos bidimensional proporcionada por la biblioteca Pandas en Python. Se utiliza para manipular y analizar datos tabulares de forma eficiente y organizada.
- **Epidemiología:** rama de la medicina que estudia la distribución, frecuencia y determinantes de las enfermedades en poblaciones humanas, con el objetivo de controlar y prevenir su propagación.
- **Simulación en Tiempo Real:** técnica de modelado que reproduce el comportamiento de un sistema dinámico en tiempo real, permitiendo una representación continua y actualizada del fenómeno bajo estudio.

- **OMS:** Siglas de la Organización Mundial de la Salud, es una agencia especializada de las Naciones Unidas, fundada en 1948, cuya misión es promover y proteger la salud a nivel global.
- **OPS:** Siglas de la Organización Panamericana de la Salud, es un organismo internacional de salud pública que actúa como la oficina regional de la OMS para las Américas.

HOJA DE APROBACIÓN

CALIFICACIÓN:

Prof. Teresa de Coronel
Secretaria General

Unidad Pedagógica San Ignacio Guazú

Mg. Esperanza del Puerto de Benítez
Directora

Unidad Pedagógica San Ignacio Guazú