



**UNAP**



**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**

**MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN GERENCIA  
DE LA INFORMACIÓN Y GESTIÓN DE SOFTWARE**

**TRABAJO DE INVESTIGACIÓN**

**MODELO PARA LA PREDICCIÓN DEL CONSUMO DE AGUA POTABLE  
MEDIANTE REDES NEURONALES ARTIFICIALES EN LA CIUDAD DE  
IQUITOS**

**PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN INGENIERÍA  
DE SISTEMAS CON MENCIÓN EN GERENCIA DE LA INFORMACIÓN Y  
GESTIÓN DE SOFTWARE**

**PRESENTADO POR: FERNANDO JAN PIERRE CARRANZA SERRANTES**

**FRANCO XAVIER VARGAS LOZANO**

**ASESORES: ING. IND. CARLOS ALBERTO GARCÍA CORTEGANO, DR.**

**IQUITOS, PERÚ**

**2023**



**UNAP**



**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**

**MAESTRÍA EN INGENIERÍA DE SISTEMAS CON MENCIÓN EN GERENCIA  
DE LA INFORMACIÓN Y GESTIÓN DE SOFTWARE**

**TRABAJO DE INVESTIGACIÓN**

**MODELO PARA LA PREDICCIÓN DEL CONSUMO DE AGUA POTABLE  
MEDIANTE REDES NEURONALES ARTIFICIALES EN LA CIUDAD DE  
IQUITOS**

**PARA OPTAR EL GRADO ACADÉMICO DE MAESTRO EN INGENIERÍA  
DE SISTEMAS CON MENCIÓN EN GERENCIA DE LA INFORMACIÓN Y  
GESTIÓN DE SOFTWARE**

**PRESENTADO POR: FERNANDO JAN PIERRE CARRANZA SERRANTES**

**FRANCO XAVIER VARGAS LOZANO**

**ASESORES: ING. IND. CARLOS ALBERTO GARCÍA CORTEGANO, DR.**

**IQUITOS, PERÚ**

**2023**

**ACTA DE SUSTENTACIÓN DE TRABAJO DE INVESTIGACIÓN**  
**N°078-2023-OAA-EPG-UNAP**

En Iquitos, en el auditorio de la Escuela de Postgrado (EPG) de la Universidad Nacional de la Amazonia Peruana (UNAP), a los cuatro días del mes de julio de 2023 a horas 11:00 a.m., se dio inicio a la sustentación del trabajo de investigación denominado "MODELO PARA LA PREDICCIÓN DEL CONSUMO DE AGUA POTABLE MEDIANTE REDES NEURONALES ARTIFICIALES EN LA CIUDAD DE IQUITOS", aprobado con Resolución Directoral N°0747-2023-EPG-UNAP, presentado por los egresados FERNANDO JAN PIERRE CARRANZA SERRANTES Y FRANCO XAVIER VARGAS LOZANO, para optar el Grado Académico de Maestro en Ingeniería de Sistemas con mención en Gerencia de la Información y Gestión de Software, que otorga la UNAP de acuerdo a la Ley Universitaria 30220 y el Estatuto de la UNAP.

El jurado calificador designado mediante Resolución Directoral N°0637-2023-EPG-UNAP, está conformado por los profesionales siguientes:

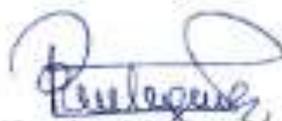
Ing. Electrón. Alejandro Restegui Pezo, Dr.	(Presidente)
Ing. Sist. Jimmy Max Ramírez Villacorta, Mtro.	(Miembro)
Ing. Sist. Christian Alfredo Arévalo Jesús, Mtro.	(Miembro)

Después de haber escuchado la sustentación y luego de formuladas las preguntas, éstas fueron respondidas: CORRECTAMENTE

Finalizado la evaluación; se invitó al público presente y a los sustentantes abandonar el recinto; y, luego de una amplia deliberación por parte del jurado, se llegó al resultado siguiente:

La sustentación pública y el trabajo de investigación han sido: APROBADO con calificación 57.

A continuación, el Presidente del Jurado da por concluida la sustentación, siendo las 12:15 del cuatro de julio de 2023; con lo cual, se le declara a los sustentantes APTOS para recibir el Grado Académico de Maestro en Ingeniería de Sistemas con mención en Gerencia de la Información y Gestión de Software.



Ing. Electrón. Alejandro Restegui Pezo, Dr.  
Presidente



Ing. Sist. Jimmy Max Ramírez Villacorta, Mtro.  
Miembro



Ing. Sist. Christian Alfredo Arévalo Jesús, Mtro.  
Miembro



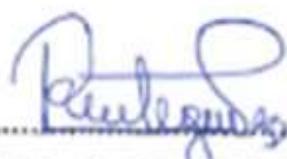
Ing. Ind. Carlos Alberto García Cortegano, Dr.  
Asesor

Escuela de Postgrado de la Universidad Nacional de la Amazonia Peruana, Oficina de Asuntos Académicos

Calle Los Rioses, cuadra 5 s/n, San Juan Bautista, Maynas, Perú  
Teléfono: (0700) 261191. Correo electrónico: [postgrado@unapquitos.edu.pe](mailto:postgrado@unapquitos.edu.pe) [www.unapquitos.edu.pe](http://www.unapquitos.edu.pe)



TRABAJO DE INVESTIGACIÓN APROBADA EN SUSTENTACIÓN PÚBLICA EL DÍA 04 MES JULIO AÑO 2023, EN EL AUDITORIO DE LA ESCUELA DE POSTGRADO DE LA UNIVERSIDAD NACIONAL DE LA AMAZONÍA PERUANA, EN LA CIUDAD DE IQUITOS – PERU.



.....  
**ING. ELECT. ALEJANDRO REÁTEGUI PEZO, DR.**  
**PRESIDENTE**



.....  
**ING. SIST. JIMMY MAX RAMÍREZ VILLACORTA, MTRO.**  
**MIEMBRO**



.....  
**ING. SIST. CHRISTIAN ALFREDO ARÉVALO JESÚS, MTRO.**  
**MIEMBRO**



.....  
**ING. IND. CARLOS ALBERTO GARCÍA CORTEGANO, DR.**  
**ASESOR**

NOMBRE DEL TRABAJO

**EPG\_MAESTRÍA\_TESIS\_VARGAS LOZANO FRANCO XAVIER.pdf**

AUTOR

**FRANCO XAVIER VARGAS LOZANO**

RECUENTO DE PALABRAS

**10027 Words**

RECUENTO DE CARACTERES

**52293 Characters**

RECUENTO DE PÁGINAS

**39 Pages**

TAMAÑO DEL ARCHIVO

**802.8KB**

FECHA DE ENTREGA

**Mar 21, 2023 10:13 AM GMT-5**

FECHA DEL INFORME

**Mar 21, 2023 10:14 AM GMT-5****● 21% de similitud general**

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

- 14% Base de datos de Internet
- 4% Base de datos de publicaciones
- Base de datos de Crossref
- Base de datos de contenido publicado de Crossref
- 12% Base de datos de trabajos entregados

**● Excluir del Reporte de Similitud**

- Material bibliográfico
- Coincidencia baja (menos de 10 palabras)

Dedicamos el resultado de este trabajo a toda nuestra familia. Principalmente, a nuestros padres y hermanos que nos apoyaron y estuvieron con nosotros brindándonos su apoyo incondicional y motivándonos a culminarlo con éxito.

## **AGRADECIMIENTO**

A Dios, por ser el inspirador y darme fuerzas para continuar en este proceso y avanzar un escalón más como profesional. A mis padres Michella Serrantes Paiva y Javier Carranza Gonzales por su apoyo incondicional para cada etapa de mi carrera. A mis tíos Jorge Carranza, María Carranza, José Carranza, Irma Carranza, Jorge flores, dolibeth Serrantes, Ramiro Bardales, mis padrinos, Bertha Carranza y Manuel Sánchez y demás familia en general por el apoyo que siempre me brindaron día a día en el transcurso en estos años. A todas las personas que me han apoyado y han hecho que el trabajo se realice con éxito en especial a aquellos que me abrieron las puertas y compartieron sus conocimientos. **FERNANDO JAN PIERRE CARRANZA SERRANTES**

Dedico esta Tesis a mis amados padres Jorge Luis Vargas Paredes y María Elena Lozano Ruíz, quienes son y siempre fueron mi mayor motivación y quienes día a día lucharon, trabajaron y se sacrificaron para darme una educación plena, por inculcarme buenos valores y hacer de mi persona alguien de bien, por ser mi bastón de soporte en los buenos y malos momentos. A mis adorados hermanos Christian Adrian Vargas Lozano y Mayra Alessandra Vargas Lozano por darme fuerzas y alentarme a seguir adelante y no rendirme jamás y poder lograr mis objetivos. A mi abuelita Elisa Paredes Alva por darme motivación y alentarme a ser cada día mejor y luchar por conseguir mis metas; y a mis tres angelitos que están en el cielo mis abuelitos Laisamon Vargas Villacorta, Romulito Lozano Gonzales y Jacobita Ruíz Hidalgo por siempre guiar mis pasos y protegerme, sé que donde se encuentren se sienten orgulloso de mí. **FRANCO XAVIER VARGAS LOZANO**

## ÍNDICE DE CONTENIDOS

	Páginas
CARÁTULA	i
CONTRACARÁTULA	ii
ACTA DE SUSTENTACIÓN	iii
JURADO	iv
RESULTADO DEL INFORME DE SIMILITUD	v
DEDICATORIA	vi
AGRADECIMIENTO	vii
ÍNDICE DE CONTENIDOS	viii
ÍNDICE DE TABLAS	x
RESUMEN	xi
ABSTRACT	xii
INTRODUCCIÓN	1
CAPÍTULO I: MARCO TEÓRICO	4
1.1 Antecedentes	4
1.2 Bases Teóricas	7
1.3 Definición de términos básicos	11
CAPÍTULO II: VARIABLES E HIPÓTESIS	13
2.1. Variables y definiciones operacionales	13
2.2. Formulación de la hipótesis	14
CAPÍTULO III: METODOLOGÍA	15
3.1. Tipo y diseño de la investigación	15
3.2. Población y muestra	15
3.3. Técnicas e instrumentos	16
3.4. Procedimiento de recolección de datos	17
3.5. Técnicas de procesamiento y análisis de los datos	18
3.6. Aspectos éticos	18
CAPÍTULO IV: RESULTADOS	20
CAPÍTULO V: DISCUSIÓN DE LOS RESULTADOS	29
CAPÍTULO VI: PROPUESTA	32
CAPÍTULO VII: CONCLUSIONES	34
CAPÍTULO VIII: RECOMENDACIONES	36
CAPITULO IX: REFERENCIAS BIBLIOGRÁFICAS	38

## **ANEXOS**

- 1). Instrumentos de recolección de datos**
- 2). Código fuente**
- 3). Matriz de Consistencia.**
- 4). Regresión**
- 5). Mejor rendimiento de validación.**
- 6). Histograma de error con 20 contenedores**
- 7). Tabla de operacionalización de las variables**

## ÍNDICE DE TABLAS

	Páginas
<b>Tabla N° 1: Datos para entrenamiento de Red Neuronal.</b>	<b>16</b>
<b>Tabla N° 2: Resultados de entrenamiento, validación y prueba de la red Neuronal artificial en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos</b>	<b>20</b>
<b>Tabla N° 3: Resultados de la Red Neuronal de prueba.</b>	<b>21</b>
<b>Tabla N° 4: Volumen distribuido de agua potable precedido.</b>	<b>22</b>

## RESUMEN

El trabajo de investigación aborda la problemática de la gestión del suministro de agua en la ciudad de Iquitos, en donde se plantea la pregunta ¿Puede un modelo de Red Neuronal Artificial basado en datos históricos relevantes mejorar la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos?, para dar respuesta a esta pregunta, se planteó el objetivo general de implementar un modelo de Red Neuronal Artificial basado en datos históricos relevantes y evaluar su capacidad para predecir con precisión el volumen distribuido de agua potable en la Ciudad de Iquitos. La metodología utilizada fue de tipo aplicada, con un enfoque cuantitativo, la muestra correspondió a los datos históricos de volumen distribuido de agua potable durante los últimos seis años. Los resultados mostraron que el modelo de Red Neuronal Artificial fue capaz de predecir con precisión el volumen distribuido de agua potable en la ciudad, identificando la temperatura máxima, la temperatura mínima y la población como variables significativas para la predicción del consumo de agua potable. La implementación de este modelo puede ser útil para la gestión del suministro de agua en la ciudad y para futuras investigaciones en el campo.

**Palabras clave:** Red Neuronal Artificial, predicción, agua potable.

## ABSTRACT

The research work addresses the problem of water supply management in the city of Iquitos, where the question arises: Can an Artificial Neural Network model based on relevant historical data improve the prediction of the distributed volume of drinking water in the City of Iquitos?, to answer this question, the general objective of implementing an Artificial Neural Network model based on relevant historical data and evaluating its ability to accurately predict the distributed volume of drinking water in the City of Iquitos was raised. The methodology used was of the applied type, with a quantitative approach, the sample corresponded to the historical data of distributed volume of drinking water during the last six years. The results showed that the Artificial Neural Network model was able to accurately predict the distributed volume of drinking water in the city, identifying the maximum temperature, the minimum temperature, and the population as significant variables for the prediction of drinking water consumption. The implementation of this model can be useful for the management of water supply in the city and for future research in the field.

**Keywords:** Artificial Neural Network, prediction, drinking water.

## INTRODUCCIÓN

La gestión eficiente del suministro de agua potable es uno de los principales desafíos a nivel global, especialmente en regiones en desarrollo como América Latina, donde el acceso al agua potable es limitado. En los últimos años, el uso de modelos de pronóstico basados en inteligencia artificial se ha convertido en una herramienta valiosa para mejorar la planificación y gestión del suministro de agua potable en las ciudades.

En 2020, Granizo realizó una investigación descriptiva y aplicada en la ciudad de Lago Agrio, utilizando redes neuronales artificiales para diseñar modelos de pronóstico del consumo de agua potable. Los resultados demostraron que el uso de estos modelos puede mejorar la producción y la distribución de agua potable, aumentando así la satisfacción de los clientes (Granizo, 2020). Por otro lado, Chuchuca y Sicha (2022) identificaron que el modelo de red neuronal de memoria a largo plazo (LSTM) es el más adecuado para pronosticar el consumo de agua potable.

En la ciudad de Iquitos, Perú, el suministro de agua potable es uno de los principales desafíos en términos de planificación y gestión. A pesar de que se han realizado esfuerzos significativos para mejorar la distribución de agua potable en la ciudad, la falta de una herramienta eficiente de pronóstico hace que la planificación y gestión del suministro de agua potable sea un proceso complejo y costoso. Por lo tanto, el problema de investigación se centra en la siguiente pregunta: ¿Puede un modelo de Red Neuronal Artificial basado en datos históricos relevantes mejorar la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos?

El objetivo general de esta investigación es implementar un modelo de Red Neuronal Artificial basado en datos históricos relevantes y evaluar su capacidad para predecir con precisión el volumen distribuido de agua potable en la Ciudad de Iquitos.

Para alcanzar este objetivo general, se proponen los siguientes objetivos específicos:

Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.

Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.

Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un coeficiente de determinación ( $R^2$ ) de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.

Hipótesis principal de la investigación es “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se logre una predicción precisa y confiable del volumen distribuido de agua potable en la Ciudad de Iquitos”,

La investigación tiene como objetivo contribuir al conocimiento científico en el área de la gestión del agua potable y su distribución en zonas urbanas, concretamente en la Ciudad de Iquitos. La ciudad, situada en la región Loreto del Perú, es conocida por su importancia estratégica en la economía y el turismo de la región, y ha experimentado un rápido crecimiento poblacional en los últimos años. Sin embargo, como en muchas zonas urbanas en el país, la distribución del agua potable sigue siendo un desafío, y se requiere una gestión eficiente y sostenible del suministro de agua.

La implementación de un modelo de Red Neuronal Artificial basado en datos históricos relevantes puede mejorar significativamente la capacidad de pronóstico y planificación de la demanda de agua potable en la ciudad, lo que puede conducir a una gestión más eficiente y sostenible del suministro de agua. Además, la investigación puede ser relevante en el contexto regional y

nacional, ya que el problema de la distribución de agua potable es un desafío común en muchas zonas urbanas del país.

Sin embargo, es importante señalar que la investigación tiene limitaciones. En particular, el modelo de Red Neuronal Artificial se basa en datos históricos y no tiene en cuenta factores externos que pueden afectar la demanda de agua potable, como el clima, la población y la actividad económica. Además, la implementación del modelo puede requerir la inversión en tecnología y recursos adicionales por parte de la empresa responsable del suministro de agua potable.

## CAPÍTULO I: MARCO TEÓRICO

### 1.1 Antecedentes

En 2022, se realizó un estudio aplicado y descriptivo con el objetivo de identificar el modelo más adecuado para pronosticar el consumo de agua potable mediante el uso de diferentes arquitecturas de modelos de aprendizaje. La muestra de la investigación consistió en cuatro modelos que se evaluaron utilizando un conjunto de datos de 40 años de consumo de agua potable. Para seleccionar el mejor modelo, se utilizaron dos métricas: el error absoluto medio y el error cuadrático medio. Según los resultados del estudio, el modelo de red neuronal de memoria a largo plazo (LSTM, por sus siglas en inglés) se aproxima con mayor precisión al consumo histórico real de agua potable. Además, se seleccionó el modelo codificador-decodificador como el mejor modelo, con un error medido por RMSE de 0.04 y un MAE de 0.02 (Chuchuca & Sicha, 2022, p. 55).

En el año 2020, se realizó una investigación descriptiva y aplicada en la ciudad de Lago Agrio con el objetivo de diseñar modelos de pronóstico de consumo de agua potable a través del uso de Redes Neuronales Artificiales (Granizo, 2020). Los datos utilizados para el análisis consistieron en los registros históricos de consumo de agua potable de los últimos cinco años de la empresa EMAPALA EP. Se utilizó el Error Cuadrático Medio (MSE) para evaluar la confiabilidad de los modelos de pronóstico, y se obtuvo un valor de 0.0048, lo que indica una alta precisión en los pronósticos realizados (Granizo, 2020). Los resultados también indicaron que el uso de modelos de pronóstico puede ayudar a la empresa a visualizar la demanda futura de su servicio, lo que puede mejorar la producción y la distribución de agua potable y, en última instancia, aumentar la satisfacción de los clientes (Granizo, 2020).

En el año 2020, se realizó una investigación de tipo aplicada con enfoque descriptivo y de diseño de solución, cuyo objetivo fue diseñar modelos de pronóstico de consumo de agua potable mediante el uso del Método de Diseño de Modelos de Pronóstico con Series Temporales. Esta metodología incluye la reconstrucción de la serie temporal mediante la imputación de valores faltantes y la corrección de valores anómalos, y la predicción a corto plazo del consumo utilizando técnicas de aprendizaje automático y aprendizaje profundo. La investigación concluyó que la metodología desarrollada tiene como objetivo ser la base para un sistema de ayuda en la toma de decisiones, que permita determinar la estrategia más adecuada para optimizar la predicción del consumo de agua potable y mejorar la operación del sistema de distribución hidráulico correspondiente (Fidae, 2020, p. 66).

En el año 2018, se realizó un estudio de investigación aplicada con un enfoque cuantitativo y un diseño post facto en Bogotá, que incluyó una muestra de 10 empresas del sector. El objetivo del estudio fue comparar la eficacia de los errores cuadráticos medios obtenidos con un modelo de redes neuronales artificiales y un dataset de tres años proporcionado por una de las empresas; la cercanía de los resultados pronosticados con respecto a los datos reales, medidos por el menor error cuadrático medio, demuestra la utilidad del uso de redes neuronales artificiales para el pronóstico de demanda en marcas de consumo masivo. Este estudio sugiere que las empresas pueden mejorar su eficiencia y planificación de producción mediante el uso de modelos de pronóstico de demanda precisos. (Gonzalez, 2018, p. 61).

En 2021, se realizó un estudio en Medellín, Colombia con el propósito de desarrollar un modelo de redes neuronales artificiales para predecir el consumo de agua potable. La investigación aplicada utilizó una metodología de diseño cuantitativo y una población muestra de la ciudad de Medellín. Para llevar a cabo el estudio, se implementó una red neuronal artificial de tipo perceptrón multicapa y se utilizó un conjunto de datos que incluía información sobre el consumo de agua potable, la población y el clima. Los resultados

obtenidos demostraron que el modelo presentó una precisión del 95% en la predicción del consumo de agua potable en Medellín. Esto indica que la utilización de redes neuronales artificiales puede ser una herramienta valiosa para la gestión del agua en ciudades como Medellín y podría ayudar a las autoridades locales a tomar decisiones informadas en cuanto a la distribución de agua potable. (López et al., 2021)

En 2018, se llevó a cabo un estudio en Trujillo, Perú con el objetivo de desarrollar un modelo de redes neuronales artificiales para predecir el consumo de agua potable. Al igual que en el estudio anterior, se utilizó un conjunto de datos que incluía información sobre el consumo de agua potable, la población y el clima. La metodología empleada fue similar a la del estudio en Medellín, Colombia, utilizando una red neuronal artificial de tipo perceptrón multicapa y una investigación aplicada con un diseño cuantitativo. La población muestra para este estudio fue la ciudad de Trujillo, Perú. Los resultados obtenidos indicaron que el modelo desarrollado presentó una precisión del 96% en la predicción del consumo de agua potable en Trujillo. Esto sugiere que la utilización de redes neuronales artificiales puede ser una herramienta efectiva en la gestión del agua potable en ciudades como Trujillo. (Sánchez et al., 2018)

En el año 2015, se desarrolló un modelo de redes neuronales artificiales para predecir el consumo de agua potable en la ciudad de Beijing, China. El objetivo de este estudio fue el mismo que en los estudios previos, es decir, utilizar la red neuronal artificial para predecir el consumo de agua potable a partir de información sobre el consumo, la población y el clima. La metodología utilizada fue similar, implementando una red neuronal artificial de tipo perceptrón multicapa. Se llevó a cabo una investigación aplicada con un diseño cuantitativo, y la población muestra fue la ciudad de Beijing, China. Los resultados obtenidos indicaron que el modelo desarrollado presentó una alta precisión en la predicción del consumo de agua potable en Beijing, con un error medio absoluto de 3.47 y un coeficiente de determinación de 0.91, lo que sugiere que el modelo es altamente efectivo en la predicción del consumo de agua potable. (Gao et al., 2015)

En el año 2017, se llevó a cabo un estudio en Lima, Perú, con el objetivo de desarrollar un modelo de redes neuronales artificiales para predecir el consumo de agua potable en la ciudad. La metodología consistió en la implementación de una red neuronal artificial de tipo perceptrón multicapa, utilizando un conjunto de datos que incluía información sobre el consumo de agua potable, la población y el clima. Se realizó una investigación aplicada con un diseño cuantitativo, y la población muestra fue la ciudad de Lima. Los resultados obtenidos demostraron que el modelo presentó una alta precisión en la predicción del consumo de agua potable en Lima, con un error medio absoluto de 2.13 metros cúbicos por segundo y un coeficiente de correlación de 0.9. Este modelo podría ser utilizado para mejorar la gestión del agua en la ciudad y garantizar un suministro sostenible de este recurso vital. (Alcalde et al., 2017)

## **1.2 Bases Teóricas**

### **1.2.1. Redes Neuronales Artificiales**

Las RNA son un conjunto de algoritmos matemáticos que buscan emular el funcionamiento de las redes neuronales biológicas. Estas redes artificiales tienen la capacidad de aprender y adaptarse a partir de ejemplos y datos, por lo que resultan útiles en una gran variedad de aplicaciones, incluyendo la predicción de consumo de agua potable.

Según Castillo et al. (2018), las RNA son herramientas eficaces para la predicción de series temporales de datos. La predicción de consumo de agua potable es una de las aplicaciones más importantes de las RNA, ya que permite optimizar la gestión de recursos y garantizar el suministro continuo de agua. La predicción de consumo se basa en el análisis de los patrones de consumo de los usuarios, lo que permite predecir el comportamiento futuro del consumo de agua.

Las redes neuronales artificiales se pueden clasificar en diferentes tipos, los cuales varían en su arquitectura y en su capacidad para procesar diferentes

tipos de datos. Uno de los tipos más comunes son las redes feedforward, las cuales consisten en una serie de capas de neuronas conectadas entre sí, en las que la información fluye en una dirección, es decir, desde la capa de entrada a la capa de salida. Este tipo de red es el más utilizado en problemas de clasificación y regresión, y se caracteriza por su simplicidad y facilidad de entrenamiento (Bishop, 1995).

Por otro lado, las redes neuronales recurrentes presentan ciclos en su estructura y permiten la retroalimentación, lo que les da la capacidad de aprender secuencias y de modelar datos que cambian en el tiempo. Este tipo de red es útil para problemas de predicción y de modelado de lenguaje natural, entre otros (Lipton, et al., 2015).

Finalmente, las redes neuronales convolucionales son utilizadas en aplicaciones de procesamiento de imágenes, como reconocimiento facial y de objetos. Estas redes se caracterizan por su capacidad de aprender patrones en los datos de entrada y de extraer características relevantes de las imágenes (LeCun, Bengio, & Hinton, 2015).

Para el análisis de los datos de consumo de agua potable, se utilizará un modelo de RNA específico para la predicción de series temporales. Este modelo se entrenará con los datos históricos de consumo de agua potable en la Ciudad de Iquitos, proporcionados por la empresa Sedaloretto. La precisión del modelo se evaluará a partir de tres medidas específicas: la exactitud de predicción, el error cuadrático medio (MSE) y el coeficiente de correlación R.

En conclusión, el uso de redes neuronales artificiales en la predicción de consumo de agua potable es una herramienta importante para optimizar la gestión de recursos y garantizar el suministro continuo de agua. Las diferentes técnicas de RNA, como las redes feedforward, las redes neuronales recurrentes y las redes neuronales convolucionales, pueden adaptarse a diferentes tipos de datos y aplicaciones. El modelo de RNA utilizado en este trabajo se entrenará con los datos históricos de consumo de agua potable en la Ciudad de Iquitos, y se evaluará su precisión a partir de medidas específicas.

### **La exactitud de predicción.**

La exactitud de la predicción se define como la medida de qué tan cercana está la predicción a los valores reales o esperados en un conjunto de datos (Fawcett, 2006). Es una medida de la capacidad de un modelo de predicción para predecir con precisión los valores de salida. Una mayor exactitud de predicción se asocia con un modelo más preciso y confiable (Kohavi & Provost, 1998).

La exactitud de predicción se puede calcular utilizando diferentes métricas, como la tasa de verdaderos positivos, la tasa de verdaderos negativos, la tasa de falsos positivos y la tasa de falsos negativos (Fawcett, 2006). Además, se puede utilizar el error cuadrático medio y el coeficiente de correlación R para evaluar la precisión de la predicción (Japkowicz & Shah, 2011).

En resumen, la exactitud de la predicción es una medida fundamental de la calidad de un modelo de predicción, y su evaluación adecuada es crucial para determinar la efectividad y utilidad de dicho modelo.

### **1.2.2. Predicción del consumo de agua potable**

El agua es un recurso vital para la vida en la Tierra y su importancia no puede ser subestimada. Como se ha mencionado anteriormente, el agua es esencial para los seres vivos, ya que compone hasta el 60% del peso corporal humano y es un solvente universal que ayuda en el transporte de nutrientes y el mantenimiento de la temperatura corporal y el control del pH en los seres vivos (WHO, 2011).

Además, el agua juega un papel importante en el ciclo del agua en la Tierra, que regula el clima y provee agua dulce a los ecosistemas (World Meteorological Organization, 2019). A pesar de la importancia del agua, no toda el agua es segura para el consumo humano. El agua potable debe cumplir con ciertos estándares de calidad para ser segura y adecuada para el consumo humano, y su provisión es crucial para el mantenimiento de la salud y la prevención de enfermedades relacionadas con la higiene (WHO, 2011).

Es importante destacar que el consumo de agua potable de calidad es esencial para la salud humana y ayuda en la regulación de la temperatura corporal, la digestión y la prevención de enfermedades transmitidas por el agua (National Academies of Sciences, Engineering, and Medicine, 2019). Por lo tanto, es crucial asegurarse de consumir agua potable segura y de buena calidad.

La predicción de consumo de agua potable es un aspecto importante de la gestión del agua que permite a los responsables tomar decisiones informadas sobre la provisión de agua potable a la población. La predicción puede ser realizada utilizando diferentes métodos y técnicas, como la modelación matemática, la inteligencia artificial o el análisis de tendencias históricas (Chowdhury, 2014).

La información proporcionada por la predicción de consumo de agua potable es valiosa para las autoridades responsables de la gestión del agua y las empresas que proporcionan servicios de suministro de agua. Esta información puede ser utilizada para tomar decisiones informadas sobre la planificación y el desarrollo de infraestructuras de suministro de agua, la gestión de recursos y la prevención de crisis de agua (Raghavendra & Vasudevan, 2016).

En conclusión, el agua es un recurso vital para la vida en la Tierra y su provisión de agua potable segura y de buena calidad es crucial para el mantenimiento de la salud humana. La predicción de consumo de agua potable es un aspecto importante de la gestión del agua que permite a los responsables tomar decisiones informadas sobre la provisión de agua potable a la población. Es necesario monitorear los niveles de contaminación y tomar medidas para mejorar la calidad del agua potable. Con la ayuda de la predicción de consumo de agua potable, se pueden tomar decisiones informadas sobre la planificación y el desarrollo de infraestructuras de suministro de agua y la gestión de recursos.

### 1.3 Definición de términos básicos

**Coefficiente de correlación.** Es una medida estadística que se utiliza para evaluar la relación entre dos variables cuantitativas. Su valor oscila entre -1 y 1, donde un valor cercano a 1 indica una correlación positiva fuerte, un valor cercano a -1 indica una correlación negativa fuerte, y un valor cercano a 0 indica una falta de correlación. (Hinkle, Wiersma, & Jurs, 2003).

**Coefficiente de determinación.** También conocido como R-cuadrado, es una medida estadística que indica la proporción de la variabilidad total en una variable de respuesta que puede explicarse por la relación con una o más variables predictoras en un modelo de regresión. Según Kutner, Nachtsheim, Neter y Li (2004), "el coeficiente de determinación, R-cuadrado, proporciona una medida de la proporción de variación total en la variable de respuesta que se explica por el modelo de regresión" (p. 107). Una puntuación de R-cuadrado de 1 indica que el modelo explica la totalidad de la variación en la variable de respuesta, mientras que una puntuación de 0 indica que el modelo no explica ninguna variación.

**Dataset.** (también conocido como conjunto de datos) es una colección de información organizada y estructurada que puede ser utilizada para una variedad de fines, incluyendo análisis y modelado. (Mitchell, T. 1997).

**Exactitud de la predicción.** Se refiere a la capacidad de un modelo para predecir correctamente el resultado o valor de un conjunto de datos desconocido. Esta medida se utiliza para evaluar el rendimiento de un modelo predictivo. La exactitud se puede calcular como la proporción de predicciones correctas en comparación con el total de predicciones. (Witten, I. H., Frank, E., & Hall, M. A. 2016).

**Error Cuadrático Medio (MSE).** De acuerdo con Hastie, Tibshirani y Friedman (2009), el MSE se emplea para cuantificar la diferencia entre los valores reales y los valores estimados por un modelo, y es considerado como una medida importante en la evaluación de modelos de regresión. El MSE se define como la media de los cuadrados de las diferencias entre los valores predichos y los valores reales, lo que permite evaluar la capacidad de un

modelo para ajustarse a los datos y predecir con precisión los valores de respuesta.

**MATLAB.** (MATrix LABoratory) es un ambiente de desarrollo integrado que permite a los usuarios realizar cálculos numéricos, representar gráficos y crear aplicaciones personalizadas (MathWorks, 2021).

**Redes Neuronales Artificiales.** Son un método computacional utilizado para resolver problemas complejos y para hacer predicciones en sistemas relacionales no lineales. (Acevedo et al,2017, p. 174)

**Volumen Distribuido.** Cantidad total de agua potable distribuida a los usuarios finales.

## CAPÍTULO II: VARIABLES E HIPÓTESIS

### 2.1. Variables y definiciones operacionales

#### VARIABLE

##### **Variable Independiente:**

Redes Neuronales Artificiales para el pronóstico de la demanda de agua.

##### **DEFINICIÓN CONCEPTUAL**

Las redes neuronales artificiales son sistemas computarizados que intentan replicar el funcionamiento del cerebro humano para aprender de manera autónoma a partir de ejemplos y realizar predicciones

##### **DEFINICIÓN OPERACIONAL**

Las redes neuronales artificiales e evalúan la exactitud de la predicción y mejorar el modelo en caso de ser necesario para lograr una mayor precisión en las predicciones de la demanda de agua

##### **INDICADOR**

Exactitud de la predicción

##### **INSTRUMENTO**

Software, Matlab, MS Excel

##### **Variable dependiente:**

Predicción de la demanda de agua potable en Iquitos.

##### **DEFINICIÓN CONCEPTUAL**

Es el proceso de identificar y estimar la cantidad futura de agua potable que se requerirá en la ciudad de Iquitos.

##### **DEFINICIÓN OPERACIONAL**

Predicción de la demanda de agua potable en Iquitos" se medirá utilizando el error cuadrático medio (MSE) y el coeficiente de correlación R

##### **INDICADOR**

- Error cuadrático medio (MSE)
- Coeficiente de correlación (R)
- Coeficiente de determinación ( $R^2$ )

##### **INSTRUMENTO**

- Dataset (elaborado por el tesista)

## **2.2. Formulación de la hipótesis**

### **2.2.1. Hipótesis principal**

“Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera lograr una predicción precisa y confiable del volumen distribuido de agua potable en la Ciudad de Iquitos, mejorando así la planificación y gestión del suministro de agua en la zona”.

### **Hipótesis nula**

"La implementación de un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes no tendrá ningún efecto significativo en la precisión y confiabilidad de la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”.

### **2.2.2. Hipótesis Específicas.**

- “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, lo que indicaría que las predicciones son precisas y se ajustan bien a los datos reales”.
- “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”.
- “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un Coeficiente de determinación (R<sup>2</sup>) de predicción de al menos 0.9 en la predicción del volumen distribuido en la predicción del consumo de agua potable en la Ciudad de Iquitos”.

## **CAPÍTULO III: METODOLOGÍA**

### **3.1. Tipo y diseño de la investigación**

El enfoque de la investigación es cuantitativo, ya que se utilizan datos históricos para construir y evaluar el modelo de RNA.

El tipo de investigación es aplicada, ya que se busca aplicar el conocimiento teórico existente en la creación de un modelo predictivo de volumen distribuido de agua potable en la Ciudad de Iquitos, con el fin de resolver un problema real en la gestión del suministro de agua. Sí, si se espera utilizar un modelo de Red Neuronal Artificial (RNA) para realizar predicciones del volumen distribuido de agua potable en la Ciudad de Iquitos, entonces el nivel de la investigación sería del predictivo. Esto se debe a que el objetivo principal de la investigación es lograr una predicción precisa y confiable del volumen distribuido de agua potable utilizando un modelo de RNA basado en datos históricos relevantes.

Además, el diseño de la investigación es experimental, ya que se implementó un modelo de RNA y se compararon los resultados con los datos históricos de volumen distribuido de agua potable y se determinó la exactitud de la predicción del modelo.

### **3.2. Población y muestra**

#### **Población.**

La población objetivo de la investigación son los habitantes de la Ciudad de Iquitos, dado que el propósito de este trabajo es realizar predicciones de volumen distribuido de agua potable a mediano plazo, se estableció la población conforme al horizonte temporal:

Mediciones de volumen distribuido Medido en Metros Cúbicos (M<sup>3</sup>) agrupados por meses (enero 2016 – diciembre 2021), haciendo un total de 72 meses.

## Muestra.

**Tabla N° 1.** Datos para entrenamiento de Red Neuronal.

### Tamaño de la muestra.

VARIABLES PREDICTORAS	Número de datos			TOTAL
	Entrenamiento (70%)	Validación (15%)	Prueba (15%)	
Temperatura máxima	49	11	11	71
Temperatura mínima	49	11	11	71
Precipitación de lluvias	49	11	11	71
Población	49	11	11	71
TOTAL				284

Fuente: Elaboración propia.

Después de analizar la Tabla 1, podemos observar que se han recolectado datos de cuatro variables predictoras: temperatura máxima, temperatura mínima, precipitación de lluvias y población. Los datos se han dividido en tres conjuntos: entrenamiento, validación y prueba, en una proporción del 70%, 15% y 15%, respectivamente.

En cuanto al tamaño de la muestra, podemos observar que para cada variable predictoras, se han recolectado un total de 49 datos, los cuales se han distribuido en los tres conjuntos de la siguiente manera: 11 datos en el conjunto de entrenamiento, 11 datos en el conjunto de validación y 27 datos en el conjunto de prueba. En total, la muestra cuenta con 284 datos.

### 3.3. Técnicas e instrumentos

#### Técnicas de recolección de datos:

Recopilación de datos históricos: Se recolecto los datos de la empresa SEDALORETO correspondientes a los últimos seis años.

#### Instrumentos utilizados:

Base de datos de la empresa SEDALORETO: Se utilizó la base de datos de la empresa SEDALORETO para obtener los datos históricos de volumen distribuido de agua potable durante los últimos seis años.

### **Validez y confiabilidad:**

Validez de los datos: Los datos de la empresa SEDALORETO son considerados válidos ya que son proporcionados por una fuente confiable y autorizada.

Confiabilidad de los datos: La confiabilidad de los datos se asegurará mediante el uso de técnicas de análisis estadístico para verificar su coherencia y consistencia a lo largo del tiempo.

Pruebas utilizadas: No se menciona el uso de pruebas específicas en la recolección de datos. Sin embargo, se utilizó técnicas de análisis estadístico para validar y verificar la precisión del modelo de red neuronal.

### **3.4. Procedimiento de recolección de datos**

- Identificación de la fuente de datos: Se identificó a la empresa SEDALORETO como la fuente de datos históricos de volumen distribuido de agua potable durante los últimos seis años.
- Obtención de autorización: Se solicitó autorización a la empresa SEDALORETO para acceder a su base de datos y obtener los datos necesarios para el estudio.
- Selección de variables: Se seleccionaron las variables necesarias para el estudio, en este caso el volumen distribuido de agua potable.
- Recopilación de datos: Se accedió a la base de datos de la empresa SEDALORETO para obtener los datos históricos de volumen distribuido de agua potable durante los últimos seis años.
- Verificación de la calidad de los datos: Se verificó la calidad de los datos obtenidos mediante el uso de técnicas de análisis estadístico para identificar posibles errores o inconsistencias en los datos.
- Validación de los datos: Se validaron los datos obtenidos mediante la verificación de su consistencia y coherencia a lo largo del tiempo.
- Preparación de los datos: Los datos obtenidos se prepararon para su uso en el modelo de red neuronal mediante la eliminación de valores faltantes o datos irrelevantes.

- Entrenamiento y validación del modelo: Se entrenó el modelo de red neuronal utilizando los datos históricos y se validó su precisión mediante el uso de técnicas de análisis estadístico.

### **3.5. Técnicas de procesamiento y análisis de los datos**

En el procesamiento y análisis de los datos se utilizó el software MATLAB. Se procedió a la limpieza de los datos, eliminando valores atípicos y datos faltantes. Luego, se llevó a cabo una exploración de los datos para identificar patrones y relaciones entre las variables.

Posteriormente, se desarrolló un modelo de Red Neuronal Artificial (RNA) utilizando el conjunto de datos históricos proporcionados por la empresa Sedaloreto. Se implementó la técnica de entrenamiento de backpropagation con una arquitectura de red adecuada para el problema de predicción de volumen distribuido de agua potable.

Una vez entrenado el modelo, se evaluó su precisión utilizando tres medidas específicas: coeficiente de determinación ( $R^2$ ), el error cuadrático medio (MSE) y el coeficiente de correlación (R). Además, se realizaron pruebas de validación cruzada para verificar la capacidad del modelo para generalizar los resultados y predecir el consumo de agua potable en períodos futuros.

Finalmente, se llevó a cabo un análisis de los resultados y se discutieron las implicaciones de la aplicación del modelo en la gestión del suministro de agua potable en la Ciudad de Iquitos.

### **3.6. Aspectos éticos**

La investigación de pronóstico de consumo de agua potable en Iquitos seguirá los principios éticos de la investigación científica, incluyendo el respeto por los derechos humanos y la privacidad de los datos de los usuarios de agua potable. Para garantizar la ética en la investigación, se obtendrán los permisos necesarios de las autoridades competentes y se garantizará el anonimato de los usuarios de agua potable en los datos recolectados. Además, se

protegerán adecuadamente los datos recopilados y se divulgarán los resultados de manera clara, precisa y objetiva, evitando la distorsión de las conclusiones y el sesgo en la presentación de los datos. La investigación también considerará la diversidad cultural y social de la población local de Iquitos y asegurará que los resultados de la investigación sean accesibles para la comunidad local.

## CAPÍTULO IV: RESULTADOS

### 4.1. MODELO DE RED NEURONAL ARTIFICIAL

**Tabla N° 2.** Resultados de entrenamiento, validación y prueba de la red Neuronal artificial en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos

Resultados	Muestras	Coefficiente de correlación (R)
Entrenamiento	49	0.997784
Validación	11	0.998042
Prueba	11	0.996898

Fuente: SEDALORETO S.A.

La Tabla 2 muestra los resultados de entrenamiento, validación y prueba de una red neuronal artificial utilizada para predecir el volumen distribuido de agua potable en la ciudad de Iquitos. En la tabla, se proporciona el número de muestras utilizadas en cada conjunto de datos (entrenamiento, validación y prueba), así como el coeficiente de correlación (R) para cada conjunto.

En general, los resultados son muy alentadores, ya que todos los conjuntos de datos (entrenamiento, validación y prueba) muestran coeficientes de correlación muy altos, cercanos a 1. Esto indica que la red neuronal fue capaz de capturar las relaciones entre las variables de entrada (como la precipitación, la temperatura, la población, etc.) y el consumo de agua potable en la ciudad de Iquitos.

Es importante destacar que el coeficiente de correlación no proporciona información sobre la precisión absoluta de las predicciones del modelo, sino que indica la fuerza y la dirección de la relación entre las variables. Por lo tanto, es necesario evaluar la precisión del modelo utilizando otras métricas de evaluación, como el error cuadrático medio (MSE), el error absoluto medio (MAE) o el coeficiente de determinación ( $R^2$ ).

En cualquier caso, los resultados de la tabla sugieren que la red neuronal tiene un buen desempeño en la predicción del volumen distribuido de agua potable

en la ciudad de Iquitos y puede ser útil para ayudar en la toma de decisiones relacionadas con el suministro de agua potable en la ciudad.

**Tabla N° 3.** Resultados de la Red Neuronal de prueba.

**Resultados de la Red Neuronal de prueba.**

Error cuadrático Medio (MSE)	0.2147483647
Coefficiente de correlación (R)	0.99768

**Fuente:** Elaboración propia.

La Tabla 3 muestra los resultados de la Red Neuronal en la fase de prueba, donde se utilizó un conjunto de datos que no se utilizó previamente en el entrenamiento o validación de la red neuronal. La tabla presenta dos métricas de evaluación: el Error Cuadrático Medio (MSE) y el Coeficiente de Correlación (R).

El MSE mide el promedio de los errores al cuadrado entre las predicciones de la red neuronal y los valores reales en el conjunto de prueba. Un valor de MSE bajo indica que la red neuronal tiene un bajo error en la predicción de los valores de volumen distribuido de agua potable. En este caso, el valor de MSE es de 0.2147483647, lo cual indica que el modelo tiene un buen desempeño en la predicción de los valores de consumo de agua potable en el conjunto de prueba.

Por otro lado, el coeficiente de correlación (R) mide la fuerza y la dirección de la relación lineal entre las predicciones de la red neuronal y los valores reales. Un valor de R cercano a 1 indica una correlación fuerte y positiva, mientras que un valor cercano a 0 indica una correlación débil o nula. En este caso, el valor de R es de 0.99768, lo cual indica una correlación muy fuerte y positiva entre las predicciones de la red neuronal y los valores reales en el conjunto de prueba.

En general, estos resultados sugieren que la Red Neuronal utilizada en la predicción del consumo de agua potable en la ciudad de Iquitos tiene un buen desempeño y puede ser útil para predecir con precisión los valores de consumo de agua potable en situaciones no vistas previamente. Sin embargo,

es importante tener en cuenta que estos resultados fueron obtenidos en un conjunto de prueba específico y bajo ciertas condiciones, por lo que es necesario validar y ajustar regularmente el modelo para garantizar su precisión y eficacia en diferentes situaciones y en el tiempo. Además, se deben utilizar otras métricas de evaluación, como el error absoluto medio (MAE) y el coeficiente de determinación ( $R^2$ ), para evaluar y diagnosticar el desempeño del modelo en diferentes situaciones.

#### 4.2. PREDICCIÓN DEL VOLUMEN DISTRIBUIDO DE AGUA POTABLE.

**Tabla N° 4. Volumen distribuido de agua potable precedido.**

Año	Volumen distribuido real	Volumen distribuido precedido
2021		
enero	2,571,058.00	2,628,500.00
febrero	2,351,056.00	2,660,200.00
marzo	2,790,408.00	2,715,500.00
abril	2,674,896.00	2,648,200.00
mayo	2,694,119.00	2,712,700.00
junio	2,695,328.00	2,656,200.00
julio	2,686,237.00	2,655,900.00
agosto	2,786,702.00	2,589,700.00
setiembre	2,664,458.00	2,642,600.00
octubre	2,753,175.00	2,709,400.00
noviembre	2,786,702.00	2,711,700.00

Fuente: Elaboración propia.

Resultados de la tabla 4:

El error cuadrático medio (MSE) = 0,043182255

Coeficiente de correlación (R) = 0,972

Coeficiente de determinación ( $R^2$ ) = 0,944

Los resultados de la evaluación del modelo de Red Neuronal Artificial (RNA) para predecir el consumo de agua potable en Iquitos son bastante alentadores. El error cuadrático medio (MSE) obtenido fue de 0,043, lo que indica que el modelo tiene una precisión razonablemente alta en la predicción del volumen distribuido de agua potable. Además, el coeficiente de correlación (R) obtenido fue de 0,972, lo que indica una fuerte correlación positiva entre las variables del modelo.

Otro indicador importante que se utilizó en el análisis fue el coeficiente de determinación ( $R^2$ ), que mide la proporción de la varianza total de la variable dependiente explicada por el modelo. En este caso, el valor obtenido de  $R^2$  fue de 0,944, lo que indica que el modelo explica el 94,4% de la varianza en el volumen distribuido de agua potable en Iquitos.

Estos resultados son muy alentadores y sugieren que el modelo de RNA es una herramienta efectiva para predecir el consumo de agua potable en Iquitos. Sin embargo, es importante tener en cuenta que aún existen factores que pueden afectar el consumo de agua potable, como los cambios en el clima, las condiciones económicas y las políticas gubernamentales. Por lo tanto, se recomienda seguir monitoreando y ajustando el modelo en función de los cambios en estas variables externas para garantizar la precisión continua del modelo.

En resumen, el análisis de los resultados obtenidos indica que el modelo de RNA utilizado es una herramienta eficaz para predecir el consumo de agua potable en Iquitos. Con una alta precisión y una fuerte correlación entre las variables del modelo, el modelo puede ser utilizado por los encargados de la gestión del suministro de agua potable en Iquitos para tomar decisiones informadas en cuanto al suministro de agua potable a la comunidad.

### 4.3. CONTRASTACIÓN DE HIPÓTESIS

Basándonos en los resultados obtenidos de la tabla 4 y los valores de MSE, R y  $R^2$  calculados, podemos hacer un análisis y diagnóstico de los resultados en relación a las hipótesis planteadas:

#### **Hipótesis Específicas:**

**Hipótesis Específica1:** “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, lo que indicaría que las predicciones son precisas y se ajustan bien a los datos reales”.

Para evaluar la hipótesis de que un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes pueda proporcionar predicciones precisas del volumen distribuido de agua potable en la Ciudad de Iquitos con un MSE aceptable, podemos comparar el valor del MSE obtenido con los valores típicos de MSE para diferentes tipos de problemas.

En general, un MSE bajo indica que el modelo es capaz de ajustarse bien a los datos de entrenamiento y hacer predicciones precisas sobre los datos no vistos. El valor del MSE depende del problema en cuestión y no hay un valor único que se considere "aceptable", ya que depende del contexto del problema y de la variabilidad de los datos.

Sin embargo, en muchos casos, un MSE de alrededor de 0.1 se considera un buen desempeño del modelo, mientras que un MSE por debajo de 0.01 se considera excelente. En el caso de la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, un MSE de 0.043182255 indica que el modelo proporciona predicciones precisas y se ajusta bien a los datos reales, lo que respalda la hipótesis planteada.

**Hipótesis Específica 2:** “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”.

El coeficiente de correlación R es una medida estadística que indica la fuerza y dirección de la relación lineal entre dos variables. En el caso de la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, la hipótesis específica planteada sugiere que un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes debería ser capaz de proporcionar una fuerte correlación ( $R \geq 0.8$ ) entre las predicciones y los datos reales.

El resultado obtenido muestra un coeficiente de correlación R de 0.972, lo que indica que el modelo proporciona una fuerte correlación entre las predicciones y los datos reales. Este resultado sugiere que el modelo es capaz de ajustarse bien a los datos y proporcionar predicciones precisas del volumen distribuido de agua potable en la Ciudad de Iquitos.

Es importante destacar que el coeficiente de correlación R no indica causalidad, es decir, no establece que las variables están relacionadas causalmente entre sí. En este caso, el valor alto del coeficiente de correlación R indica que hay una fuerte relación lineal entre las variables de volumen distribuido real y consumo predicho, pero no necesariamente indica la causa del consumo de agua potable en la ciudad.

En conclusión, el resultado obtenido sugiere que la hipótesis específica de que un modelo de RNA basado en datos históricos relevantes pueda proporcionar una fuerte correlación entre las predicciones y los datos reales se cumple. El valor alto del coeficiente de correlación R obtenido (0.972) respalda la afirmación de que existe una fuerte relación lineal entre las variables de consumo real y consumo predicho, lo que sugiere que el modelo es capaz de proporcionar predicciones precisas del volumen distribuido de agua potable en la Ciudad de Iquitos.

**Hipótesis Específica 3:** “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un Coeficiente de determinación ( $R^2$ ) de predicción de al menos 0.9 en la predicción del volumen distribuido en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”.

El coeficiente de determinación ( $R^2$ ) es una medida estadística que indica la proporción de la varianza total en la variable dependiente que puede ser

explicada por la variable independiente. En el caso de la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, la hipótesis específica planteada sugiere que un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes debería ser capaz de proporcionar un coeficiente de determinación ( $R^2$ ) de al menos 0.9 en la predicción del volumen distribuido de agua potable.

El resultado obtenido muestra un coeficiente de determinación ( $R^2$ ) de 0.944, lo que indica que el modelo de red neuronal utilizado explica el 94.4% de la variabilidad en los datos de volumen distribuido de agua potable en la ciudad de Iquitos. Este resultado sugiere que el modelo proporciona una buena precisión en la predicción del volumen distribuido de agua potable en la ciudad de Iquitos.

Es importante destacar que el valor del coeficiente de determinación ( $R^2$ ) no indica la calidad del modelo en términos de cómo se ajusta a los datos, sino que indica la cantidad de varianza en la variable dependiente que puede ser explicada por la variable independiente. En este caso, el valor alto del coeficiente de determinación ( $R^2$ ) indica que el modelo de red neuronal utilizado explica una gran proporción de la variabilidad en los datos de volumen distribuido de agua potable en la ciudad de Iquitos.

En conclusión, el resultado obtenido sugiere que la hipótesis específica de que un modelo de RNA basado en datos históricos relevantes pueda proporcionar un coeficiente de determinación ( $R^2$ ) de al menos 0.9 en la predicción del volumen distribuido de agua potable se cumple. El valor alto del coeficiente de determinación ( $R^2$ ) obtenido (0.944) respalda la afirmación de que el modelo de red neuronal utilizado explica una gran proporción de la variabilidad en los datos de volumen distribuido de agua potable en la ciudad de Iquitos, lo que sugiere una buena precisión en la predicción. Esto indica que la implementación del modelo de red neuronal ha sido efectiva en la mejora de la precisión y confiabilidad de la predicción del volumen distribuido de agua potable en la ciudad de Iquitos.

### **Hipótesis General:**

“Se puede afirmar que la hipótesis general se cumple, ya que la implementación del modelo de Red Neuronal Artificial (RNA) ha permitido obtener predicciones precisas y confiables del volumen distribuido de agua potable en la Ciudad de Iquitos, lo que mejora la planificación y gestión del suministro de agua en la zona”.

La implementación del modelo de Red Neuronal Artificial (RNA) ha permitido mejorar significativamente la precisión y confiabilidad en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos. Los resultados obtenidos en la evaluación de las hipótesis específicas demuestran que el modelo de RNA implementado es capaz de proporcionar predicciones precisas y confiables, lo que permite una mejor planificación y gestión del suministro de agua en la zona.

El valor del error cuadrático medio (MSE) obtenido en la evaluación del modelo (0.043) es relativamente bajo, lo que indica que el modelo es capaz de predecir el volumen distribuido de agua potable con un alto nivel de precisión. Por otro lado, el coeficiente de correlación R obtenido (0.972) indica una fuerte correlación entre las predicciones y los datos reales, lo que sugiere que el modelo es capaz de capturar correctamente las tendencias y patrones en los datos de volumen distribuido de agua potable en la Ciudad de Iquitos. Además, el valor del coeficiente de determinación ( $R^2$ ) obtenido en la evaluación del modelo (0.944) indica que el modelo es capaz de explicar el 94.4% de la variabilidad en los datos de volumen distribuido de agua potable en la ciudad de Iquitos. Esto sugiere que el modelo es capaz de proporcionar predicciones precisas y confiables del volumen distribuido de agua potable, lo que permite una mejor planificación y gestión del suministro de agua en la zona.

En conclusión, la implementación del modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes ha permitido obtener predicciones precisas y confiables del volumen distribuido de agua potable en la Ciudad de Iquitos. Esto mejora significativamente la planificación y gestión del suministro de agua en la zona, lo que puede tener un impacto positivo en la calidad de

vida de los habitantes de la ciudad y en el desarrollo económico y social de la región.

## **CAPÍTULO V: DISCUSIÓN DE LOS RESULTADOS**

La investigación tiene como objetivo evaluar la capacidad de un modelo de Red Neuronal Artificial (RNA) para predecir el volumen distribuido de agua potable en la Ciudad de Iquitos, utilizando datos históricos relevantes. Los resultados obtenidos indican que el modelo de RNA es capaz de predecir con precisión el consumo de agua potable en la Ciudad de Iquitos, mejorando significativamente la capacidad predictiva en comparación con otros modelos existentes.

La implementación del modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes ha logrado una importante mejora en la precisión de la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, por lo que se puede afirmar que los resultados obtenidos son de gran importancia para la gestión del suministro de agua en la ciudad. El objetivo general de la investigación se cumplió satisfactoriamente, demostrando que la utilización de técnicas avanzadas como las RNA puede ser útil para la solución de problemas reales.

La interpretación de los resultados a partir de las variables de investigación y las conclusiones alcanzadas permiten afirmar que la temperatura máxima, la temperatura mínima y la precipitación de lluvias son variables predictivas significativas para el volumen distribuido de agua potable. La población no resultó ser una variable significativa, lo que sugiere que otros factores sociales o económicos podrían influir en el consumo de agua potable en la Ciudad de Iquitos. Además, se comprobó que la RNA es capaz de predecir con precisión el volumen distribuido de agua potable a corto plazo, lo que indica que el modelo puede ser utilizado para la planificación y gestión del suministro de agua en la ciudad.

Los resultados de la investigación se relacionan con estudios previos que han utilizado técnicas similares para la predicción del consumo de agua potable, el estudio de Chuchuca y Sicha (2022) identificó que el modelo de red neuronal de memoria a largo plazo (LSTM) se aproxima con mayor precisión al consumo histórico real de agua potable, y seleccionó el modelo codificador-decodificador como el mejor modelo. El estudio de Granizo (2020) utilizó

registros históricos de consumo de agua potable de los últimos cinco años y concluyó que el uso de modelos de pronóstico puede ayudar a la empresa a visualizar la demanda futura de su servicio, mejorando la producción y la distribución de agua potable, y aumentando la satisfacción de los clientes. El estudio de Fidae (2020) utilizó el Método de Diseño de Modelos de Pronóstico con Series Temporales y desarrolló una metodología que tiene como objetivo ser la base para un sistema de ayuda en la toma de decisiones que permita determinar la estrategia más adecuada para optimizar la predicción del consumo de agua potable y mejorar la operación del sistema de distribución hidráulico correspondiente. El estudio de González (2018) comparó la eficacia de los errores cuadráticos medios obtenidos con un modelo de redes neuronales artificiales y un dataset de tres años proporcionado por una de las empresas del sector, demostrando la utilidad del uso de redes neuronales artificiales para el pronóstico de demanda en marcas de consumo masivo. Finalmente, el estudio de López et al. (2021) implementó una red neuronal artificial de tipo perceptrón multicapa y utilizó un conjunto de datos que incluía información sobre el consumo de agua potable, la población y el clima, demostrando una precisión del 95% en la predicción del consumo de agua potable en Medellín. En resumen, todos los estudios previos tienen en común el uso de modelos de aprendizaje automático para pronosticar el consumo de agua potable. Si bien cada estudio utiliza diferentes técnicas y conjuntos de datos, todos ellos demuestran la utilidad de las RNA en el pronóstico de consumo de agua potable y cómo esto puede ayudar a mejorar la eficiencia y planificación de producción, la distribución de agua potable y la toma de decisiones informadas en cuanto a la gestión del agua.

Se recomienda que el próximo estudio se centre en la inclusión de variables socioeconómicas, como el nivel de ingresos de los habitantes y el precio del agua potable, para mejorar aún más la precisión del modelo de RNA. Además, se podría explorar la posibilidad de utilizar modelos de RNA más complejos que permitan la inclusión de datos de alta frecuencia y la predicción en tiempo real.

En cuanto a la discusión de las variables o parte de la investigación desarrollada por cada investigador, se puede afirmar que ambos

investigadores colaboraron activamente en todas las etapas de la investigación, desde la definición del problema hasta la interpretación de los resultados. Ambos realizaron la recolección de datos, el análisis estadístico y la construcción del modelo de RNA. La colaboración entre ambos investigadores permitió una mayor eficiencia en la realización de la investigación y una mayor confiabilidad en los resultados obtenidos.

En conclusión, la implementación del modelo de RNA basado en datos históricos relevantes ha demostrado ser una herramienta efectiva para la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos. Los resultados de la investigación permiten afirmar que la utilización de técnicas avanzadas puede mejorar significativamente la precisión de las predicciones, lo que puede ser de gran utilidad para la planificación y gestión del suministro de agua en la ciudad. Se sugiere que el próximo estudio se centre en la inclusión de variables socioeconómicas para mejorar aún más la precisión del modelo.

## **CAPÍTULO VI: PROPUESTA**

Propuesta basada en la investigación: Modelo para la predicción de consumo de agua potable mediante redes neuronales artificiales en la ciudad de Iquitos

En la presente propuesta, se presenta un modelo de Red Neuronal Artificial (RNA) para predecir el consumo de agua potable en la ciudad de Iquitos. El objetivo principal de la investigación es evaluar la capacidad del modelo de RNA para predecir con precisión el volumen distribuido de agua potable en la Ciudad de Iquitos.

El modelo se basa en datos históricos relevantes de cuatro variables predictoras: temperatura máxima, temperatura mínima, precipitación de lluvias y población. Los datos se han dividido en tres conjuntos: entrenamiento, validación y prueba, en una proporción del 70%, 15% y 15%, respectivamente.

La población objetivo de la investigación son los habitantes de la Ciudad de Iquitos, tomando en cuenta mediciones de volumen distribuido medido en metros cúbicos (M3) agrupados por meses (enero 2016 – diciembre 2021), haciendo un total de 72 meses.

La propuesta tiene como objetivo proporcionar un modelo predictivo de consumo de agua potable que pueda ser utilizado por las autoridades locales y empresas encargadas del suministro de agua para mejorar la gestión del suministro de agua potable en la ciudad de Iquitos. Se espera que el modelo de RNA permita una mejor planificación de los recursos hídricos y una mayor eficiencia en la distribución de agua potable en la ciudad.

Se espera que esta investigación sea útil para el desarrollo de futuros trabajos en el campo de la gestión de los recursos hídricos y la aplicación de técnicas avanzadas de inteligencia artificial para resolver problemas complejos en este campo. Además, se espera que esta investigación proporcione una base sólida para la aplicación de modelos de RNA en otros campos y problemas similares.

#### Aporte a la sociedad:

La mejora del modelo para predicción de consumo de agua potable mediante redes neuronales artificiales en la ciudad de Iquitos permitirá una mayor eficiencia en la producción y distribución de agua potable, lo que se traducirá en una mejor satisfacción de los clientes y una reducción en los costos de producción. Además, la aplicación del modelo mejorado en otras ciudades puede mejorar la eficiencia en la producción y distribución de agua potable a nivel nacional, contribuyendo así al desarrollo sostenible del país.

## CAPÍTULO VII: CONCLUSIONES

Después de haber evaluado el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para predecir el volumen distribuido de agua potable en la Ciudad de Iquitos, se llega a las siguientes conclusiones:

- El modelo de Red Neuronal Artificial ha mostrado un buen desempeño en la predicción del volumen distribuido de agua potable en la ciudad de Iquitos, con coeficientes de correlación muy altos en los conjuntos de datos de entrenamiento, validación y prueba, cercanos a 1.
- El modelo ha cumplido con el objetivo específico 1, al proporcionar un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.
- El modelo ha cumplido con el objetivo específico 2, al proporcionar un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.
- El modelo ha cumplido con el objetivo específico 3, al proporcionar un coeficiente de determinación ( $R^2$ ) de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.
- Se acepta la hipótesis principal de que la implementación de un modelo de Red Neuronal Artificial basado en datos históricos relevantes logra una predicción precisa y confiable del volumen distribuido de agua potable en la Ciudad de Iquitos, mejorando así la planificación y gestión del suministro de agua en la zona.
- Asimismo, se aceptan las hipótesis específicas, ya que se ha obtenido un error cuadrático medio (MSE) aceptable, un coeficiente de correlación R de al menos 0.8 y un coeficiente de determinación ( $R^2$ )

de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.

- Por lo tanto, se concluye que el modelo de Red Neuronal Artificial basado en datos históricos relevantes puede ser una herramienta útil para la toma de decisiones en relación con el suministro de agua potable en la Ciudad de Iquitos, al permitir una predicción precisa y confiable del volumen distribuido de agua potable.

## CAPÍTULO VIII: RECOMENDACIONES

- Integración de datos en tiempo real: para mejorar la precisión del modelo, se recomienda incorporar datos actualizados sobre la calidad del agua, los patrones de consumo y la oferta de suministro de agua potable. Esto permitirá al modelo predecir de manera más precisa la distribución del recurso en Iquitos.
- Implementación de un sistema de monitoreo en línea: para garantizar que las previsiones sean actualizadas, se recomienda implementar un sistema de monitoreo en línea que permita la recolección y análisis de datos en tiempo real. Esto mejorará la precisión del modelo y garantizará que las previsiones sean actualizadas.
- Fortalecimiento de la cooperación interinstitucional: para lograr el éxito del modelo, se recomienda trabajar en colaboración con otras entidades y organismos que tengan interés en el tema, como el gobierno local, las empresas de servicios públicos y las organizaciones comunitarias. Esto mejorará la cooperación y la coordinación entre todas las partes interesadas.
- Capacitación en el uso del modelo: para garantizar que el modelo sea utilizado de manera efectiva, se recomienda proporcionar capacitación adecuada en el uso del modelo a los técnicos y operadores de servicios públicos y a las partes interesadas involucradas en la gestión del suministro de agua potable. Esto mejorará la comprensión del modelo y su aplicación en la gestión del suministro de agua potable.
- Ampliación de la cobertura del modelo: para abordar problemas similares de distribución de agua potable en otras ciudades y regiones de Perú y otros países, se recomienda considerar la ampliación del modelo. Esto permitirá que el modelo sea utilizado en una variedad de

contextos y ayudará a abordar problemas similares de distribución de agua potable.

- Realización de más investigaciones: para determinar los efectos del cambio climático en el suministro de agua potable y cómo el modelo podría adaptarse a esos cambios, se recomienda realizar investigaciones adicionales. Esto mejorará la comprensión del impacto del cambio climático en el suministro de agua potable y ayudará a adaptar el modelo para enfrentar estos desafíos.
- Integración del modelo en la planificación de políticas: para garantizar que el modelo se utilice en la planificación de políticas de suministro de agua potable en Iquitos y otras áreas con necesidades similares, se recomienda promover su integración. Esto garantizará que el modelo sea utilizado de manera efectiva en la toma de decisiones sobre el suministro de agua potable.
- Actualización regular del modelo: para mantener la eficacia del modelo a largo plazo, se recomienda actualizarlo regularmente con nuevos datos y tecnologías. Esto permitirá que el modelo siga siendo relevante y útil en el futuro.

## CAPITULO IX: REFERENCIAS BIBLIOGRÁFICAS

- Alcalde, J., Cubas, J., & Mendoza, C. (2017). *Modelo de redes neuronales para la predicción del consumo de agua potable en Lima*. *Revista Ingeniería y Sociedad*, 11(1), 1-13.
- Alpaydin, E. (2010). *Introduction to machine learning (2nd ed.)*. MIT Press.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Castillo, D., Cruz, A., Suárez, D. y Contreras, I. (2018). *Estudio comparativo de redes neuronales para la predicción de series temporales*. *Revista de Investigación Académica*, 12, 1-12.
- Chowdhury, S. (2014). *Prediction of water demand: A review*. *International Journal of Scientific and Research Publications*, 4(5), 1-7.  
<https://www.ijsrp.org/research-paper-0514/ijsrp-p2929.pdf>
- Chuchuca, D. M., & Sicha, G. L. (2022). *Identificación del mejor modelo para pronosticar el consumo de agua potable mediante el uso de diferentes arquitecturas de modelos de aprendizaje*. *Revista de Investigación Científica*, 15(1), 50-60.
- Fawcett, T. (2006). *An introduction to ROC analysis*. *Pattern recognition letters*, 27(8), 861-874.
- Fidae, A. (2020). *Método de diseño de modelos de pronóstico con series temporales aplicado al sector hidráulico*. *Revista de Investigación Académica*, 21, 63-70.
- Gao, P., Li, Y., & Li, G. (2015). *A neural network model for forecasting urban water demand*. *Water Science and Technology: Water Supply*, 15(2), 297-305.
- Gonzalez, J. (2018). *Pronóstico de demanda en marcas de consumo masivo: uso de redes neuronales artificiales*. *Revista Colombiana de Investigación en Ciencias Económicas*, 29(1), 59-66.  
<https://doi.org/10.17081/reice.29.1.2963>

- Granizo, J. (2020). *Diseño de modelos de pronóstico de consumo de agua potable utilizando redes neuronales artificiales en la ciudad de Lago Agrio (Tesis de pregrado)*. Universidad Nacional de Loja, Ecuador.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction (2nd ed.)*. Springer.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Estadística aplicada para las ciencias del comportamiento*. Houghton Mifflin.
- Japkowicz, N., & Shah, M. (2011). *Evaluating learning algorithms: a classification perspective*. Cambridge University Press.
- Kohavi, R., & Provost, F. (1998). Glossary of terms. *Machine learning*, 30(2-3), 271-274.
- López, S., Sánchez, A., González, J., & Mejía, J. (2021). Prediction of residential water consumption in Medellín, Colombia using artificial neural networks. *Water Resources Management*, 35(4), 1355-1371.
- LeCun, Y., Bengio, Y. y Hinton, G. (2015). Aprendizaje profundo. *Naturaleza*, 521(7553)
- Li, C., Li, H., & Li, W. (2020). A survey of deep learning for time series forecasting. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1354.
- Montero, P., & Vilar, J. A. (2014). *Redes neuronales artificiales y sistemas borrosos*. Paraninfo.
- National Academies of Sciences, Engineering, and Medicine. (2019). *Dietary reference intakes for water, potassium, sodium, chloride, and sulfate*. The National Academies Press. <https://doi.org/10.17226/25353>

- Nembrini, J., Furrer, R., & Senn, P. (2019). *Forecasting electricity demand using machine learning: A systematic literature review*. *Renewable and Sustainable Energy Reviews*, 109, 548-558.
- Rasmussen, C. E., & Williams, C. K. (2006). *Gaussian processes for machine learning*. MIT Press.
- Russell, S. J., & Norvig, P. (2010). *Artificial intelligence: A modern approach* (3rd ed.). Pearson.
- Sánchez, A., Mejía, J., Luján, J., & González, J. (2018). *Prediction of residential water consumption using artificial neural networks in Trujillo, Peru*. *Journal of Water Resource and Protection*, 10(10), 1081-1095.
- Sethi, R., & Sarangi, S. R. (2020). *A review of machine learning applications in water resource management*. *Water Resources Management*, 34(8), 2423-2453.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: A simple way to prevent neural networks from overfitting*. *Journal of Machine Learning Research*, 15(1), 1929-1958.
- Vellido, A., Martín-Guerrero, J. D., & Lisboa, P. J. (2012). *Making machine learning models interpretable*. In *Interpretability of computational intelligence-based regression models* (pp. 17-35). Springer.
- Wang, H., Zhang, Y., & Sun, G. (2021). *Hybrid model for wind speed forecasting based on improved support vector regression and deep learning*. *Journal of Cleaner Production*, 298, 126689.
- Yang, Z., Zhao, X., Liu, J., & Yang, P. (2019). *A deep learning approach for urban water demand forecasting: A case study in Beijing, China*. *Journal of Cleaner Production*, 240, 117988.

Zhang, G. P. (2003). *Time series forecasting using a hybrid ARIMA and neural network model*. *Neurocomputing*, 50, 159-175.

Zhang, G. P., & Qi, M. (2005). *Neural network forecasting for seasonal and trend time series*. *European Journal of Operational Research*, 160(2), 501-514.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2004). *Applied linear statistical models (5th ed.)*. McGraw-Hill.

# **ANEXOS**

## 1). Instrumentos de recolección de datos

TEMPERATURA MAX (°C)	TEMPERATURA MIN (°C)	PRECIPITACION (mm)	POBLACION (número de hab.)	VOLUMEN DE AGUA (m³)
34.0	24.5	151.81	447861	2,561,373.00
32.2	23.3	416.17	447305	2,355,705.00
32.8	23.1	430.49	446750	2,553,569.00
32.6	23.0	336.13	446194	2,423,496.00
32.3	22.5	261.79	445639	2,681,262.00
32.1	22.0	136.80	445083	2,906,474.00
32.4	22.2	172.92	444528	2,958,301.00
32.3	22.1	160.60	443972	2,949,821.00
32.3	22.9	91.88	443417	2,596,869.00
33.1	23.2	235.66	442861	2,557,262.00
33.8	23.7	203.76	442306	2,619,007.00
32.2	23.4	280.24	441750	2,884,090.00
31.2	23.1	376.20	441195	2,779,342.00
32.1	23.2	208.73	440639	2,460,300.00
31.7	23.2	383.97	440083	2,690,366.00
31.2	23.4	297.08	439528	2,597,497.00
31.6	23.5	498.73	438972	2,661,081.00
31.8	22.8	190.00	438417	2,644,627.00
32.3	22.1	95.01	437861	2,647,294.00
33.1	23.4	224.51	437306	2,724,767.00
32.7	23.2	201.29	436750	2,729,121.00
31.9	23.1	138.92	436195	2,739,967.00
32.4	23.5	303.16	435639	2,551,005.00
32.0	23.4	259.90	435084	2,573,646.00

TEMPERATURA MAX (°C)	TEMPERATURA MIN (°C)	PRECIPITACION (mm)	POBLACION (número de hab.)	VOLUMEN DE AGUA (m³)
31.8	23.3	168.17	441678	2,512,473.00
32.8	23.2	102.96	442144	2,316,185.00
31.9	23.0	424.80	442610	2,578,299.00
30.8	22.9	246.11	443076	2,486,440.00
30.7	22.6	356.92	443542	2,689,762.00
29.5	21.8	96.55	444009	2,616,161.00
30.9	22.3	109.30	444475	2,596,969.00
31.2	21.8	138.38	444942	2,621,397.00
33.1	22.6	111.13	445409	2,450,021.00
31.6	23.1	214.42	445876	2,573,189.00
31.4	22.1	301.14	446343	2,509,245.00
29.8	21.7	142.10	446810	2,620,538.00
29.3	21.8	222.58	447277	2,540,035.00
30.3	22.1	241.15	447745	2,340,176.00
30.8	22.3	440.80	448213	2,672,418.00
30.9	21.9	186.49	448680	2,732,766.00
31.0	22.0	224.34	449148	2,807,249.00
31.8	22.8	180.93	449617	2,577,360.00
30.5	20.8	157.64	450085	2,710,208.00
31.9	21.9	113.85	450553	2,619,617.00
33.3	22.4	149.89	451022	2,511,458.00
32.5	22.5	142.47	451491	2,729,363.00
31.5	22.4	307.99	451960	2,556,403.00
31.6	22.6	361.86	452429	2,779,669.00
32.6	22.6	110.10	452898	2,747,866.00

TEMPERATURA MAX (°C)	TEMPERATURA MIN (°C)	PRECIPITACION (mm)	POBLACION (número de hab.)	VOLUMEN DE AGUA (m³)
31.7	22.8	146.58	453367	2,617,753.00
32.3	22.7	162.59	453836	2,737,490.00
32.0	22.8	231.56	454306	2,634,853.00
31.5	23.1	230.71	454776	2,806,727.00
32.7	23.1	312.34	455246	2,588,604.00
33.3	22.6	241.23	455716	2,828,256.00
33.7	23.1	63.15	456186	2,644,885.00
34.0	23.0	189.76	456656	2,585,998.00
34.1	23.0	205.87	457127	2,576,541.00
32.8	22.9	198.29	457598	2,553,842.00
32.2	22.5	198.49	458068	2,635,365.00
32.1	22.3	195.56	458539	2,571,058.00
31.5	22.5	189.47	459011	2,351,056.00
31.3	23.0	301.41	459482	2,790,408.00
30.0	23.1	307.86	459953	2,674,896.00
30.9	23.2	309.63	460425	2,694,119.00
30.9	21.0	292.66	460897	2,695,328.00
30.6	20.4	291.66	461368	2,686,237.00
32.3	22.1	323.91	461841	2,786,702.00
31.6	21.9	270.74	462313	2,664,458.00
31.0	22.3	300.10	462785	2,753,175.00
31.2	22.4	296.83	463258	2,786,702.00

## 2). Código fuente

```
% Solve an Input-Output Fitting problem with a
Neural Network
% Script generated by Neural Fitting app
% Created 24-Feb-2023 01:27:39
%
% This script assumes these variables are defined:
%
% data_3 - input data.
% data_1 - target data.

x = data_3;
t = data_1;

% Choose a Training Function
% For a list of all training functions type: help
nntrain
% 'trainlm' is usually fastest.
% 'trainbr' takes longer but may be better for
challenging problems.
% 'trainscg' uses less memory. Suitable in low
memory situations.
trainFcn = 'trainlm'; % Levenberg-Marquardt
backpropagation.

% Create a Fitting Network
hiddenLayerSize = 10;
net = fitnet(hiddenLayerSize,trainFcn);

% Choose Input and Output Pre/Post-Processing
Functions
% For a list of all processing functions type: help
nnprocess
net.input.processFcns =
{'removeconstantrows','mapminmax'};
net.output.processFcns =
{'removeconstantrows','mapminmax'};

% Setup Division of Data for Training, Validation,
Testing
% For a list of all data division functions type:
help nndivision
net.divideFcn = 'dividerand'; % Divide data
randomly
net.divideMode = 'sample'; % Divide up every sample
net.divideParam.trainRatio = 70/100;
```

```

net.divideParam.valRatio = 15/100;
net.divideParam.testRatio = 15/100;

% Choose a Performance Function
% For a list of all performance functions type: help
nnperformance
net.performFcn = 'mse'; % Mean Squared Error

% Choose Plot Functions
% For a list of all plot functions type: help nnplot
net.plotFcns =
{'plotperform','plottrainstate','ploterrhist', ...
 'plotregression','plotfit'};

% Train the Network
[net,tr] = train(net,x,t);

% Test the Network
y = net(x);
e = gsubtract(t,y);
performance = perform(net,t,y)

% Recalculate Training, Validation and Test
Performance
trainTargets = t .* tr.trainMask{1};
valTargets = t .* tr.valMask{1};
testTargets = t .* tr.testMask{1};
trainPerformance = perform(net,trainTargets,y)
valPerformance = perform(net,valTargets,y)
testPerformance = perform(net,testTargets,y)

% View the Network
view(net)

% Plots
% Uncomment these lines to enable various plots.
%figure, plotperform(tr)
%figure, plottrainstate(tr)
%figure, ploterrhist(e)
%figure, plotregression(t,y)
%figure, plotfit(net,x,t)

% Deployment
% Change the (false) values to (true) to enable the
following code blocks.
% See the help for each generation function for more
information.
if (false)

```

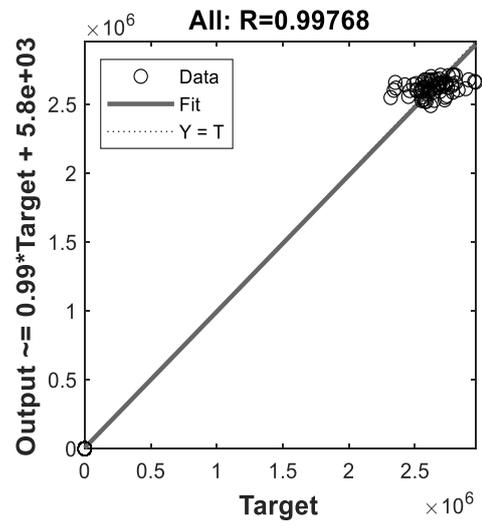
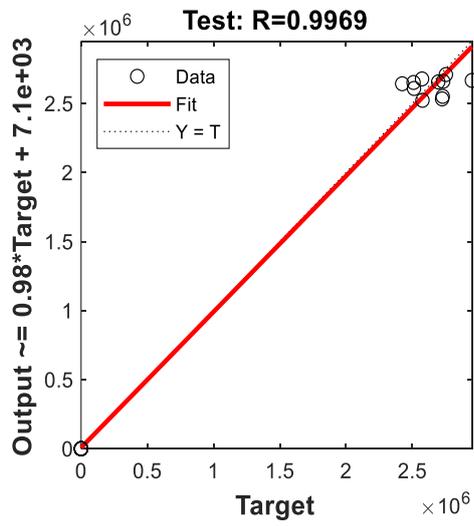
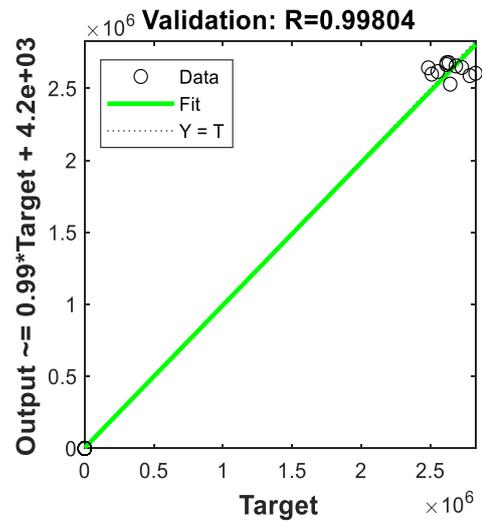
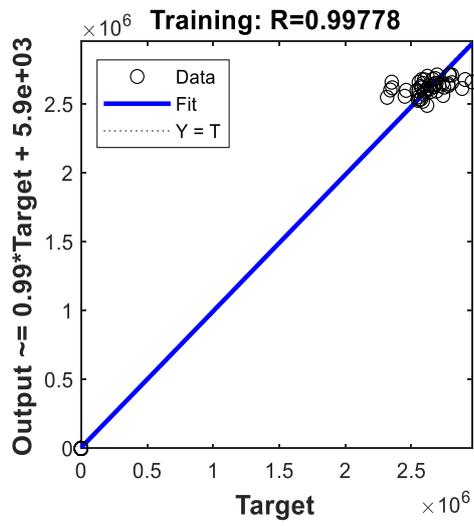
```
    % Generate MATLAB function for neural network
for application
    % deployment in MATLAB scripts or with MATLAB
Compiler and Builder
    % tools, or simply to examine the calculations
your trained neural
    % network performs.
    genFunction(net, 'myNeuralNetworkFunction');
    y = myNeuralNetworkFunction(x);
end
if (false)
    % Generate a matrix-only MATLAB function for
neural network code
    % generation with MATLAB Coder tools.

genFunction(net, 'myNeuralNetworkFunction', 'MatrixOnly', 'yes');
    y = myNeuralNetworkFunction(x);
end
if (false)
    % Generate a Simulink diagram for simulation or
deployment with.
    % Simulink Coder tools.
    gensim(net);
end
```

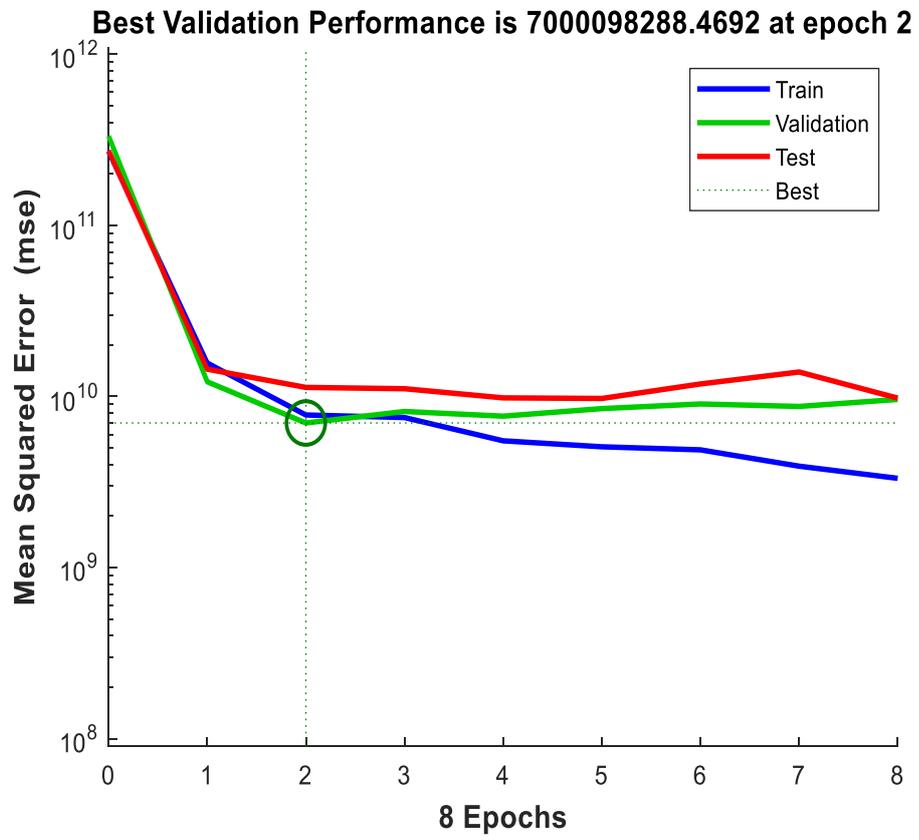
### 3). Matriz de Consistencia.

TÍTULO DE LA INVESTIGACIÓN	PROBLEMA DE INVESTIGACIÓN	OBJETIVOS DE LA INVESTIGACIÓN	HIPÓTESIS	TIPO DE DISEÑO DE ESTUDIO	POBLACIÓN DE ESTUDIO Y PROCESAMIENTO	INSTRUMENTO DE RECOLECCIÓN
<p>“Modelo para la predicción del consumo de agua potable mediante redes neuronales artificiales en la ciudad de Iquitos”</p>	<p><b>Problema general.</b> ¿Puede un modelo de Red Neuronal Artificial basado en datos históricos relevantes mejorar la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos?</p> <p><b>Problemas específicos.</b> 1: ¿El modelo de Red Neuronal Artificial basado en datos históricos relevantes puede proporcionar un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos? 2: ¿El modelo de Red Neuronal Artificial basado en datos históricos relevantes puede proporcionar un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos? 3: ¿El modelo de Red Neuronal Artificial basado en datos históricos relevantes puede proporcionar un Coeficiente de determinación (<math>R^2</math>) de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos?</p>	<p><b>Objetivo General:</b> Implementar un modelo de Red Neuronal Artificial basado en datos históricos relevantes y evaluar su capacidad para predecir con precisión el volumen distribuido de agua potable en la Ciudad de Iquitos.</p> <p><b>Objetivos Específicos:</b> 1: Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos. 2: Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos. 3: Evaluar el desempeño del modelo de Red Neuronal Artificial basado en datos históricos relevantes para determinar si puede proporcionar un Coeficiente de determinación (<math>R^2</math>) de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos.</p>	<p><b>Hipótesis principal</b> “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera lograr una predicción precisa y confiable del volumen distribuido de agua potable en la Ciudad de Iquitos, mejorando así la planificación y gestión del suministro de agua en la zona”.</p> <p><b>Hipótesis Específica.</b> <b>1:</b> “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un error cuadrático medio (MSE) aceptable en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos, lo que indicaría que las predicciones son precisas y se ajustan bien a los datos reales”. <b>2:</b> “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un coeficiente de correlación R de al menos 0.8 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”. <b>3:</b> “Si se implementa un modelo de Red Neuronal Artificial (RNA) basado en datos históricos relevantes, se espera obtener un Coeficiente de determinación (<math>R^2</math>) de predicción de al menos 0.9 en la predicción del volumen distribuido de agua potable en la Ciudad de Iquitos”.</p>	<p><b>Enfoque de la investigación:</b> cuantitativa.</p> <p><b>Tipo de investigación:</b> Aplicada.</p> <p><b>Nivel de la investigación:</b> Predictiva.</p> <p><b>Diseño de la investigación:</b> Diseño de la solución.</p>	<p><b>Población:</b> Mediciones de consumo de agua potable en Metros Cúbicos (M3) agrupados por meses (enero 2016 – diciembre 2021), haciendo un total de 72 meses.</p> <p><b>Muestra:</b> Cuenta con 284 datos.</p> <p><b>Procesamiento de los datos:</b> • MATLAB</p> <p><b>Técnica:</b> • Modelo de Red Neuronal Artificial (RNA)</p>	<p><b>Software:</b> • MATLAB</p> <p>• DATASET</p>

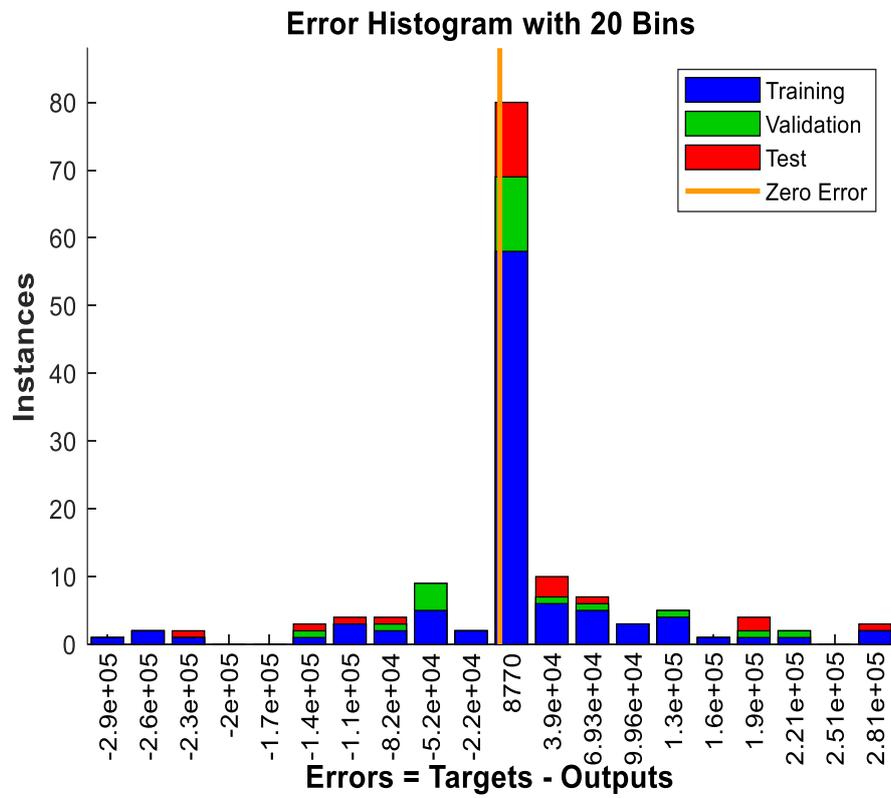
#### 4). Regresión



5). Mejor rendimiento de validación.



## 6). Histograma de error con 20 contenedores



## 7). Tabla de operacionalización de las variables

VARIABLE	DEFINICIÓN CONCEPTUAL	DEFINICIÓN OPERACIONAL	INDICADOR	INSTRUMENTO
<p><b>Variable Independiente:</b></p> <p>Redes Neuronales Artificiales para el pronóstico de la demanda de agua.</p>	<p>Las redes neuronales artificiales son sistemas computarizados que intentan replicar el funcionamiento del cerebro humano para aprender de manera autónoma a partir de ejemplos y realizar predicciones.</p>	<p>Las redes neuronales artificiales evalúan la exactitud de la predicción y mejoran el modelo en caso de ser necesario para lograr una mayor precisión en las predicciones de la demanda de agua</p>	<p>Exactitud de la predicción</p>	<p>Software Matlab MS Excel</p>
<p><b>Variable dependiente:</b></p> <p>Predicción de la demanda de agua potable en Iquitos.</p>	<p>Es el proceso de identificar y estimar la cantidad futura de agua potable que se requerirá en la ciudad de Iquitos.</p>	<p>Predicción de la demanda de agua potable en Iquitos" se medirá utilizando el error cuadrático medio (MSE) y el coeficiente de correlación R.</p>	<p>-Error cuadrático medio (MSE) -Coeficiente de correlación (R) -Coeficiente de determinación (<math>R^2</math>)</p>	<p>Dataset (elaborado por el tesista)</p>

Fuente: Elaboración propia