1. Introducción
   1. Antecedentes
   2. Justificación
   3. Hipótesis
   4. Objetivos
2. Bases Teóricas
   1. Conceptos Básicos de Series de Tiempo
   2. Conceptos básicos de Machine Learning
   3. Conceptos básicos de Redes Neuronales
      1. RNN
      2. LSTM
3. Estado del Arte
   1. Predicción de demanda
   2. Predicción de reservas de agua
   3. Sistemas complejos
4. Metodología
   1. Caso de Estudio
   2. Base de datos
   3. Modelos de predicción
   4. Evaluación de modelos
5. Resultados

FORMATO NOTAS:  
Desarrollar o citar  
Copy/paste de articulo por desmenuzar  
Contenido por revisar

Capítulo 1.

* 1. Introducción

La importancia de abordar los desafíos asociados con la gestión de agua en México es cada vez mayor. Se ha demostrado que la alteración que genera el cambio climático en los procesos ciclo hidrológicos, ha impactado a los recursos hídricos y su aprovechamiento, la agricultura, el abastecimiento poblacional, la seguridad alimentaria, y en la generación de energía hidroeléctrica [2]. Continuamente se han presentado crisis hídricas en metrópolis de México a lo largo de las últimas décadas, y en los años recientes se han visto con mayor frecuencia. La presión sobre los sistemas de agua va en aumento debido a múltiples factores tanto ambientales como sociales.

México, a través de su Plan 2030, busca alcanzar un desarrollo sostenible promoviendo políticas y medidas concretas, integradas y específicas para poner fin a la guerra contra la naturaleza [9], pero el rápido crecimiento de la población, la urbanización y las crecientes necesidades en materia de agua de los sectores agrícola, industrial y energético están provocando un aumento de la demanda de agua [6]. Informe de los Objetivos de Desarrollo Sostenible (2023) indica que, en trayectoria de las metas evaluadas en materia de Agua Limpia y Saneamiento, nos encontramos en “progreso razonable, pero es necesario acelerarlo” y “estancamiento o retroceso” [9].

La creciente frecuencia de fenómenos climáticos extremos, como las sequías prolongadas, complica la gestión de recursos hídricos en un país donde la demanda de agua aumenta debido al crecimiento demográfico y la urbanización. Los cambios en las precipitaciones y la temperatura afectarán directamente a las reservas de agua terrestre [1]. En 2018 se determinó que el 5% de aguas subterráneas presentó problemas de salinización del suelo, y que el 18% está siendo sobreexplotado [10]. Las sequías han mostrado tendencia en la última década se presentar más intensas, en años recientes el 2021 fue particular con 71% de sequías severas [10], y se extendió a poco más del 80% comenzando desde febrero [8].

Los sistemas de agua urbanos enfrentan diversas presiones derivadas de factores ambientales y sociales que pueden afectar su sostenibilidad y eficiencia. Aquí hay una descripción general de algunos de estos factores:

En términos generales, la gestión del agua se sustenta en un conjunto de políticas, atribuciones y responsabilidades mediante las cuales el Estado, los sectores económicos (principalmente agropecuario, industrial, y energético) y la sociedad en general deben promover coordinadamente el control y manejo de este recurso [10]. José Antonio Rodríguez Tirado, consultor internacional en manejo de recursos hídricos, afirma que "debería considerarse que el 'Día cero' es ahora, porque los ríos están contaminados, las fuentes están sobreexplotadas, así lo debemos de entender”. Tamara Luengo Schreck, fundadora y directora de la organización “Aqueducto”, consideró que “la crisis hídrica no es una crisis de recursos, sino más bien una crisis de gestión” [2g].

Para asistir en la gestión efectiva del agua, el presente trabajo se centra en la creación de un modelo predictivo del suministro de agua superficial, basado en redes neuronales. Este enfoque innovador busca proporcionar una herramienta avanzada y adaptable para anticipar patrones complejos y dinámicos relacionados con la gestión de los recursos hídricos. A través de la implementación de un modelo de redes neuronales, se aspira a no solo comprender las complejas interacciones entre la oferta y demanda del agua, sino también a ofrecer una contribución significativa a la toma de decisiones informada y la planificación estratégica en el ámbito hídrico. La flexibilidad de los modelos de redes neuronales promete ser adaptable a diferentes metrópolis del país y aportar una perspectiva dinámica, permitiendo una gestión más eficiente y sostenible de uno de los recursos más vitales para la vida y el desarrollo humano.

* 1. Antecedentes

México ha enfrentado desafíos significativos relacionados con la sequía y el desabasto de agua, fenómenos que han impactado tanto a zonas urbanas como rurales. Múltiples factores como disminución en las precipitaciones, crecimiento urbano, concesiones de agua y el agotamiento de acuíferos ha resultado continuamente en aguda escasez de agua. (citar)

En años más recientes, la situación se ha agravado debido a la variabilidad meteorológica y el cambio climático. La región norte de México, en particular, ha experimentado periodos prolongados de sequía, afectando la disponibilidad de agua para el consumo humano, la agricultura y la industria. Este problema se ha visto exacerbado por la sobreexplotación de acuíferos, la falta de infraestructuras adecuadas y la gestión ineficiente de los recursos hídricos. (citar)

Las zonas metropolitanas son las que más sufren de escasez debido a que el tamaño de población y actividad socioeconómicas sobrepasan los límites de infraestructura [3]. Las zonas metropolitanas con mayor número de habitantes en México son el Valle de México, la zona metropolitana de Monterrey y de Guadalajara, las cuales han padecido de escasez de agua los últimos años y actualmente [3].

El Área Metropolitana de Monterrey (AMM) en Nuevo León, México, es un ejemplo representativo del consumo urbano de agua en una región con alto estrés hídrico. Servicios de Agua y Drenaje de Monterrey (SADM), el organismo que administra el agua potable y saneamiento es considerado un modelo de gestión urbana. La dotación de agua es de 251 litros por habitante al día, excediendo el valor recomendado por la OMS de 100 litros por día para minimizar riesgos de salud. La eficiencia física es del 71.7%, lo que significa que el 28.3% del agua se pierde debido a fugas, tomas clandestinas y dotaciones extraordinarias. Aunque no es posible reducir la pérdida a 0%, los estándares internacionales sugieren que una pérdida aceptable no debería superar el 15%, y estudios en México indican que es factible mantenerla por debajo del 20% [1g]

La escasez de agua y la necesidad de implementar políticas integrales de gestión del agua no son problemas únicos de México. El pronóstico de suministro de agua se sustenta en investigaciones previas que han explorado diversas metodologías para la predicción de flujos hídricos y niveles de agua en embalses. Por ejemplo, un estudio relevante realizado para la presa “Bang Lang” aplica un modelo H2O deep learning, que consiste en un multi-layer feed-forward Artificial Neural Network (ANN) para crear un modelo predictivo de entrada de agua [1]. Además, investigaciones como la relacionada con la presa de “Three Georges” han aplicado el modelo ARIMA para prever flujos de agua, enfatizando la utilidad de métodos estadísticos en la planificación y gestión eficiente de recursos hídricos [2].

Asimismo, se ha explorado el uso de redes neuronales artificiales para la predicción del lanzamiento de agua en la presa “Pa Sak Jolasid”, demostrando la eficacia de este enfoque sobre otras técnicas exploradas como “support machine vector”, esto se debe a la capacidad de ANN para aprender modelos de datos no lineares [3]. Otro estudio hizo uso de datos sobre niveles de agua, consumo y precipitación para desarrollar 5 variantes de modelos autorregresivos convencionales como ARIMA para pronosticar los niveles de agua en la presa de Estambul, resaltando la relevancia de estos modelos en la anticipación de cambios en los niveles de agua [4].

Adicionalmente, se observa la integración de técnicas de aprendizaje profundo y aprendizaje convencional (Peak-oriented Forecasting by Model Switching) en un estudio que combinó ambos enfoques para prever los flujos de entrada en embalses, subrayando la sinergia de estos métodos en la particular tarea de predecir picos en el comportamiento del flujo del agua [5].

Estas investigaciones resaltan la diversidad de enfoques utilizados en casos similares, brindando una base sólida para la exploración de un modelo de redes neuronales que pueda abordar de manera efectiva los desafíos asociados con el pronóstico del suministro de agua en presas.

* 1. Justificación

La construcción de un modelo predictivo actualizado y capaz de ajustarse a las nuevas tendencias climatológicas es primordial para satisfacer la necesidad imperante de gestionar eficientemente los recursos hídricos. Es de mayor interés garantizar el abastecimiento sostenible de agua a la población y a las actividades económicas.

La planificación a largo plazo es esencial para una gestión efectiva del agua, considerando la variabilidad climática y las fluctuaciones en la demanda. Un modelo predictivo se presenta como una herramienta clave, permitiendo anticipar patrones estacionales y planificar inversiones en infraestructuras de manera informada. Asimismo, la optimización de recursos se vuelve imperativa ante la limitada disponibilidad de agua, y un modelo predictivo facilita la comprensión del equilibrio entre la demanda y la oferta, posibilitando la programación de suministros y la implementación de medidas de conservación. Ante la realidad del cambio climático y la previsión de eventos extremos, la creación de un modelo predictivo se erige como una herramienta esencial para adaptarse de forma dinámica a estas condiciones cambiantes. En última instancia, contar con un modelo confiable de pronóstico brinda eficiencia en la toma de decisiones, permitiendo a las autoridades asignar recursos de manera informada, implementar políticas de conservación y planificar infraestructuras resilientes frente a la creciente demanda y las condiciones climáticas en evolución constante.

Las evaluaciones en el subsector hidráulico han estado limitadas a verificar el cumplimiento de objetivos de programas nacionales y regionales, sin una evaluación profunda de la gestión del agua. Es importante que instituciones gubernamentales y universitarias incluyan en sus programas líneas de investigación para mejorar la gestión y administración del agua. Estas investigaciones deben basarse en proyectos específicos que permitan desarrollar futuros programas regionales y sectoriales más eficaces. [3g]

* 1. Hipótesis

Se plantea que un modelo de redes neuronales aplicado para predecir la demanda de agua en un sistema complejo será capaz de ofrecer pronósticos confiables a 5 años y adaptarse de manera dinámica a los cambios climatológicos y de infraestructura del sistema de agua.

Dicho modelo será adecuado al Área Metropolitana de Monterrey, pero podrá ser adecuado a otras metrópolis de México.

* 1. Objetivos

Crear un modelo predictivo utilizando Redes Neuronales para sistemas de agua para detectar bajos niveles en presas y sequías.

* Obtención de datos por parte de Servicios de Agua y Drenaje de Monterrey.
* Formación de base de datos.
* Realizar una clasificación de las presas según su comportamiento.
* Valorar la variable de agua no contabilizada mediante algoritmos de machine learning.
* Construir la red neuronal el software Python.
* Comparar resultados con datos actuales del Área Metropolitana de Monterrey.

**2. Bases Teóricas (DESARROLLAR y CITAR Y AGREGAR TEORIA Y FORMULAS)**

Se examinan tres técnicas prominentes para la predicción de series de tiempo univariadas y multivariadas: los modelos Autorregresivos Integrados de Media Móvil (ARIMA), las Redes Neuronales Recurrentes (RNN) y las Memorias a Largo Plazo (LSTM). A continuación, se detallan cada uno de estos modelos y su aplicación en la predicción de series temporales.

**2.1 Modelos Autorregresivos Integrados de Media Móvil (ARIMA) y (ARMA)**

Los modelos ARIMA son técnicas estadísticas clásicas utilizadas para el análisis y predicción de series temporales. Este enfoque combina tres componentes principales:

Autorregresión (AR): Utiliza la relación entre un valor actual y sus valores pasados.

Media Móvil (MA): Modela el error de predicción en función de errores pasados.

Uso en Series Temporales

Los modelos ARIMA son eficaces para predecir series temporales que presentan patrones lineales y estacionales. Se utilizan ampliamente en aplicaciones económicas y financieras, como la previsión de ventas, el análisis de mercado y la estimación de demanda. Su principal ventaja radica en su capacidad para manejar datos estacionales y no estacionarios a través de la diferenciación.

**2.2 Redes Neuronales Recurrentes (RNN)**

Las Redes Neuronales Recurrentes (RNN) son una clase de redes neuronales diseñadas para procesar datos secuenciales. A diferencia de las redes neuronales tradicionales, las RNN tienen conexiones que permiten la persistencia de información a través del tiempo, lo que las hace ideales para manejar datos temporales.

Las RNN son especialmente útiles para predecir series temporales donde la secuencia de datos y el contexto temporal son cruciales. Se aplican en áreas como el procesamiento del lenguaje natural, el reconocimiento de voz y la predicción de comportamiento de usuarios. Sin embargo, su capacidad para aprender dependencias a largo plazo puede ser limitada debido al problema del gradiente de desaparición.

**2.3 Redes Neuronales Convolucionales (CNN)**

**2.4 Long Short-Term Memory (LSTM)**

Las Long Short-Term Memory (LSTM) son una mejora de las RNN estándar, diseñadas para superar las limitaciones de estas en cuanto a la retención de información a largo plazo. Las LSTM introducen un mecanismo de memoria más sofisticado mediante el uso de celdas y puertas que regulan el flujo de información.

Las LSTM son extremadamente útiles para predecir series temporales con dependencias a largo plazo y dinámicas complejas. Se utilizan en aplicaciones como la predicción de series financieras, el modelado de secuencias biológicas y el análisis de señales. Su capacidad para manejar datos secuenciales con dependencias prolongadas las hace adecuadas para problemas donde las RNN estándar fallan.

En resumen, los modelos ARIMA, RNN y LSTM representan enfoques distintos para la predicción de series temporales, cada uno con sus propias ventajas y limitaciones. Los ARIMA son más adecuados para datos lineales y estacionarios, las RNN son efectivas para secuencias con dependencias a corto plazo, y las LSTM sobresalen en el manejo de dependencias a largo plazo en series temporales complejas. La elección del modelo depende de la naturaleza de los datos y de los objetivos específicos de la predicción.

**3. Estado del Arte (REFERENCIAS)(FOLDER “Supply\_Demand\_Forecast”)**

**DESCRIBIR LOS CASOS ESPECIFICOS REVISADOS PARA ESTA SECCION**

La gestión del agua es un tema de relevancia mundial, pero los problemas específicos varían en cada región, lo que requiere enfoques de predicción adaptados a las condiciones locales. En algunas áreas, puede ser crucial predecir el inflow y outflow de los sistemas hídricos debido al riesgo de inundación, mientras que, en otras, la atención se centra en la oferta y demanda de agua o en el volumen almacenado en embalses con intención de hacer una gestión efectiva del agua. Cada contexto presenta desafíos únicos, desde la disponibilidad y calidad del agua hasta la infraestructura y el clima, lo que exige soluciones de gestión y modelos predictivos específicos para cada situación.

[NOTA] El paper con mejores resultados es el [2] donde se construye una red neuronal con una combinación de capas CNN-LSTM-AM para tratar con el problema de largas secuencias con CNN y el problema de memoria con las neuronas del tipo LSTM y poder encontrar dependencias de largo tiempo lo cual es necesario cuando se tienen periodicidades de un año debido al clima. La parte AM aun no la entiendo, pero por ahora mi modelo a construir aspira a ser CNN LSTM.

**3.1 Modelos Econométricos**

Los modelos más sencillos son aplicables en la predicción de la demanda de agua debido a que esta variable suele presentar una tendencia constante a lo largo del tiempo. Estos modelos, como los ARIMA, son útiles porque pueden capturar patrones temporales y tendencias lineales en los datos históricos de demanda. Aunque simplificados, proporcionan una base sólida para anticipar el consumo futuro de agua, ayudando a los gestores a planificar y tomar decisiones informadas sobre la administración de los recursos hídricos. [¿?]

Los modelos ARIMA pueden resultar insuficientes al tratar con sistemas más complejos debido a que simplifican excesivamente la realidad y no capturan la complejidad de los fenómenos hídricos, dependen de supuestos poco realistas, y su precisión puede verse afectada por la baja calidad de los datos [7]. Además, estos modelos no se adaptan bien a cambios rápidos en las condiciones ambientales, lo que puede llevar a resultados engañosos y decisiones ineficaces en la gestión del agua [7].

In this paper, we present a comparative study for water demand forecasting using support vector linear regression and AutoRegressive Integrated Moving Average (ARIMA). The study has been carried out on the state of Kuwait daily water consumption. The result shows that ARIMA has MAPE (1.8) and RMSE (9.4) while support vector linear regression has MAPE (0.52) and RMSE (2.59) which indicates the deviation of the forecasted water demand versus the actual water consumption. [1]

ARIMA Based Forecasting of stream flows of Three Georges Dam for efficient Water Resource Planning and Management [11]

The models were based on conventional auto-regressive models (AR), moving average models (MA), auto-regressive moving average (ARMA), and ARMA with Exogenous variables (ARIMAX). [13]

Análisis econométrico de la demanda de agua en Tijuana, México Econometric analysis of water demand in Tijuana, Mexico [15]

**3.2 SVM**

Currently, support vector machine (SVM) (Ji et al. (2014); Wang et al. (2015)), autoregressive moving average model (ARMA) and autoregressive integrated moving average model (ARIMA) (Colak et al. (2015)) are commonly employed to forecast short-term water supply. [4] These traditional methods are capable of achieving tolerable results on small-size datasets. However, there are still some aspects to be considered in these methods, such as low accuracy, poor robustness and generalization. [4]

**3.3 Redes Neuronales (casos de estudio RNN, CNN, LSTM, y compuestos)**

Finally, the constructed CNN-LSTM-AM model captures the spatiotemporal information of the water supply data and makes an accurate prediction. [2]

Real-time and accurate water supply forecast is crucial for water plant. most existing methods are likely affected by factors such as weather and holidays, which lead to a decline in the reliability of water supply prediction. In this paper, we address a generic artificial neural network, called Initialized Attention Residual Network (IARN), which is combined with an attention module and residual modules. [4]

This study aims to compare the application of deep learning and conventional machine learning algorithms for predicting daily reservoir inflow. Long short-term memory (LSTM) has been applied as a deep learning algorithm and boosted regression tree (BRT) has been implemented as a machine learning algorithm. [5]

Prediction and analysis of water resources demand in Taiyuan City based on principal component analysis and BP neural network [9]

This study applied the H2O deep learning model, which is a multi-layer feed-forward Artificial Neural Network (ANN) to create the model for water inflow forecasting. [10]

Estimated for satisfactory results by an ANN precision model R2 approximately 0.88513 RMSE approximately 8.044 and the SVR model R2 result. Approx. 0.43369 RMSE Approx. 18.484. [12]

**3.4 Sistemas Multivariados de Series de Tiempo (folder “CNN\_LSTM”)**

[NOTA] En esta sección se ven casos de estudio fuera del área de gestión hídrica donde se utilizan múltiples series de tiempo, como temperatura y precipitación, para aportar información.

NDVI Forecasting Model Based on the Combination of Time Series Decomposition and CNN – LSTM [1]

A CNN–LSTM model for gold price time-series forecasting [2]

Operation Energy Consumption Estimation Method of Electric Bus Based on CNN Time Series Prediction [3]

Multivariate CNN-LSTM Model for Multiple Parallel Financial Time-Series Prediction [4]

Multiple time‑series convolutional neural network for fault detection and diagnosis and empirical study in semiconductor manufacturing [5]

**OTROS METODOS**

Results validate that the ELRF algorithm can effectively recognize sensitive features compared with decision tree and regular random forest algorithms. The model used in this study shows a strong self-adaptive ability in the modeling process of multiple regression. [3]

Implementation of supervised intelligence committee machine method for monthly water level prediction [6]

Review on applications of artificial intelligence methods for dam and reservoir-hydro-environment models [7] (es una comparative de metodos)

To forecast the peak inflow more accurately, we propose a method called Peak-oriented Forecasting by Model Switching (PForMS) which is a hybrid approach that utilizes the forecasting capabilities of both conventional machine learning and Deep Learning-based methods. [14]