

Sistema de clasificación de agua residual gris domestica empleando Redes neuronales para determinar el nivel de uso del agua en la ciudad de Durango

Reyes Ibarra Luis Alberto, Herrera Castillo Perla Alejandra

Universidad Politécnica de Durango Carr. Dgo-Mex Km 9.5 S/N Loc. Dolores Hidalgo
C.P. 34300 Durango, Dgo Tel. 618 1501323
Ing. en Redes y Telecomunicaciones
luis.reyes@unipolidgo.edu.mx, perla.herrera@unipolidgo.edu.mx

RESUMEN

Este trabajo presenta un algoritmo basado en una Red Neuronal multicapa tipo Perceptrón para clasificar el nivel de uso del agua para un proceso de tratamiento de agua gris doméstica, el algoritmo es capaz de clasificar el agua de acuerdo a las características del nivel de pH, total de sólidos disueltos y la temperatura, dicho algoritmo se implementa en un sistema embebido Raspberry para ello el modelo de red neuronal utiliza funciones de conversión de TensorFlow Lite. Una de las ventajas de implementar un modelo de clasificación inteligente es que puede dar mayor confiabilidad y certeza en el uso del agua gris tratada.

Palabras clave: Tratamiento de Aguas residual gris, Red neuronal Perceptrón, Raspberry, TensorFlow

Abstract

This work presents an algorithm based on a Perceptron-type multilayer Neural Network to classify the level of water use for a domestic gray water treatment process. The algorithm is capable of classifying water according to the characteristics of the pH level, total of dissolved solids and temperature, said algorithm is implemented in a Raspberry embedded system. For this, the neural network model uses TensorFlow Lite conversion functions. One of the advantages of implementing an intelligent classification model is that it can provide greater reliability and certainty in the use of treated gray water.

Keywords: Gray wastewater treatment, Perceptron neural network, Raspberry, TensorFlow.

1. INTRODUCCIÓN

La reutilización de agua tratada para uso doméstico es casi nula en los hogares de la ciudad de Durango, es por ello que se pretende diseñar un sistema capaz de tratar el agua residual gris que se genera principalmente de los ciclos de lavado de ropa.

El sistema que se propone cuenta con sistemas clásicos para el tratamiento de agua residual en conjunto con un sistema inteligente que emplea redes neuronales que es la parte que se trata en el presente artículo, todo ello, para asegurar el tipo de uso que se le puede dar al agua tratada como lo es para el riego de patios, utilización en depósitos de baño o riego en plantas de ornato. Con la implementación del sistema los principales beneficiados es la sociedad en general ya que contribuye a la protección del medio ambiente ya que se tendrán menores

consumos de agua al reutilizarla en diversos usos como riego de jardines, áreas verdes, lavado de coches entre otros usos.

El manejo de aguas residuales genera beneficios sociales, ambientales y económicos esenciales para el desarrollo sostenible de cualquier región. La reutilización del agua se torna viable y atractiva desde el punto de vista económico cuando permite la recuperación de gastos. Esto se logra al tratar las aguas residuales hasta alcanzar un nivel de calidad aceptable para los usuarios.

La problemática del agua en México es muy grande en cuanto a calidad y cantidad es por ello que se ha incrementado la preocupación por desarrollar sistemas que ayuden a tratar el agua residual de una manera más eficiente y con tecnologías de punta.

En la ciudad de Durango, México, no es la excepción y la problemática se agrava aún más ya que se tiene el problema de concentraciones altas de arsénico y flúor debido a que la extracción se realiza de pozos más profundos construidos en el acuífero de la Ciudad de Durango. También la cantidad de agua es escasa debido a que el acuífero Valle del Guadiana que abastece a esta ciudad está sobreexplotado y se encuentra en veda de explotación, según lo publicado en el Diario oficial de la Federación [1]

Por lo anterior es de vital importancia que las aguas residuales domésticas sean tratadas, por ello el propósito de esta investigación es diseñar un sistema de tratamiento de agua residual en específico de aguas grises domésticas empleando redes neuronales. Se combinarán etapas clásicas para el tratamiento de agua residual y sistemas de sensado, además de actuación electrónica, todo ello, para obtener un sistema capaz de clasificar el agua que fue tratada, mediante el uso de un modelo de red neuronal hasta alcanzar un nivel de calidad aceptable para los usuarios.

1.1 Problemática del agua en la ciudad de Durango

En la ciudad de Durango sus fuentes de abasto principal son las aguas subterráneas y los sistemas de captación de agua, ambas tienen problemáticas mayores donde es posible tener un desabasto mayor en los próximos cinco años de no atenderse dichas problemáticas.

Los acuíferos se ven claramente afectados por las sequías que en este año 2024 han sido extremas para la ciudad de Durango según los reportado por CONAGUA [2], la variación e

intensidad de lluvias ha sido muy poca con respecto a años anteriores lo que provoca que haya poco flujo de agua hacia el subsuelo, provoca también la sobreexplotación de pozos de agua profunda, lo que conlleva a tener agua a mayores costos y con menor calidad debido a los minerales presentes en aguas profundas. En la figura se muestra las zonas de México con mayor sequía, en Durango en los recientes años sufre de sequías extremas y severas.

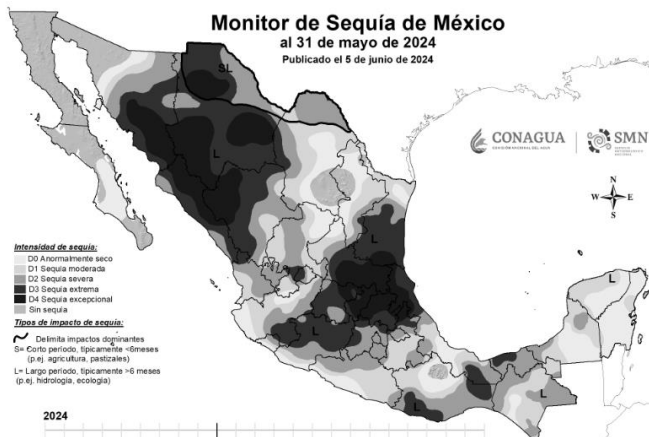


Figura 1. Monitor de sequía en México

Los sistemas de captación de agua de donde se abastece la ciudad de Durango son 3 presas las cuales en Junio de 2024 tienen un porcentaje del 40%, 60% y 11% de llenado, son niveles muy bajos para el primer semestre del año, estos datos son obtenidos de CONAGUA del monitor de presas en México [3]. En la figura 2 se muestra la tabla 1 con detalles de una de las presas principales que abastecen la ciudad de Durango.

1.2 Aguas grises

Las aguas grises son aguas que provienen de las lavadoras, regaderas, fregaderos y lavabos, las cuales tuvieron un uso ligero y pueden contener residuos como jabón, cabello, suciedad o bacterias, pero que se les puede dar un uso una vez que son tratadas [5]. Las aguas provenientes del inodoro son clasificadas como aguas negras y también algunas aguas de lavabos o tarjas que reciben productos químicos, la regulación de aguas grises y negras dentro de un hogar depende mucho de la legislación que tenga cada país.

1.3 Red neuronal multicapa tipo Perceptrón

Una red neuronal tipo Perceptrón es un tipo de red neuronal con capas densamente conectadas, contiene una capa de entrada, una capa de salida y un conjunto de capas ocultas (intermedias), dependiendo del número de conexiones que presente la red puede estar total o parcialmente conectada (ver figura 2) [7].

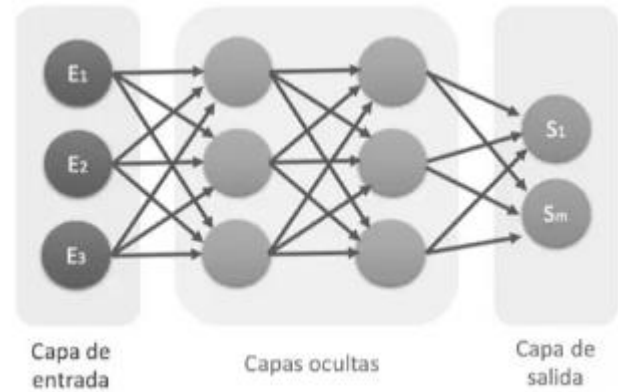


Figura 2. Red neuronal multicapa

Una forma de dar solución a un sistema de clasificación es empleando Redes neuronales capaces de realizar las tareas de clasificación. El Perceptrón Multicapa es muy utilizado para clasificar, otros algoritmos son vecinos cercanos, bosques aleatorios y árboles de decisión.

El Perceptrón multicapa para clasificación es el algoritmo elegido para el desarrollo del presente proyecto.

En el Perceptrón multicapa en su fase de aprendizaje se realizan dos pasos, la propagación hacia delante para obtener el resultado de la red y la propagación hacia atrás (backpropagation) con el objetivo de modificar los pesos de las conexiones para que el valor de la red se asemeje más al valor real, esta aproximación se realiza mediante la función gradiente del error [7].

Una red neuronal Perceptrón multicapa se emplea para dar solución a problemas de clasificación multiclase, no solo binaria 1 y 0, sino más de dos posibles clasificaciones [8].

Los Perceptrones multicapa están formados por unidades funcionales llamadas Perceptrones, en la ecuación 1 se define el Perceptrón.

$$Z = \vec{w} \cdot X + b \quad \text{Ec.1}$$

Z=Salida del Perceptrón

$\vec{w}$ = Vector de peso

X = Matriz de características

b = Inclinación

Cuando los Perceptrones se apilan, forman estructuras llamadas capas densas que luego pueden conectarse para construir una red neuronal. La ecuación de una capa densa es similar a la de un Perceptrón, pero en su lugar utiliza una matriz de peso y un vector de polarización, en la ecuación 2 se define este tipo de red neuronal.

$$Z = W \cdot X + \vec{b} \quad \text{Ec.2}$$

Z=Salida de capa densa

W = Matriz de peso

X = Matriz de características

$\vec{b}$ = Vector de sesgo

Al agregar funciones de activación no lineales a las salidas de capas densas puede ayudar al clasificador MLP a aprender límites de decisión complejos [8].

2. DESARROLLO

2.1 Diagrama general

En la figura 3 se muestra el diagrama general de este proyecto, cabe resaltar que en este artículo solo se describe el desarrollo del sistema de clasificación empleando redes neuronales.

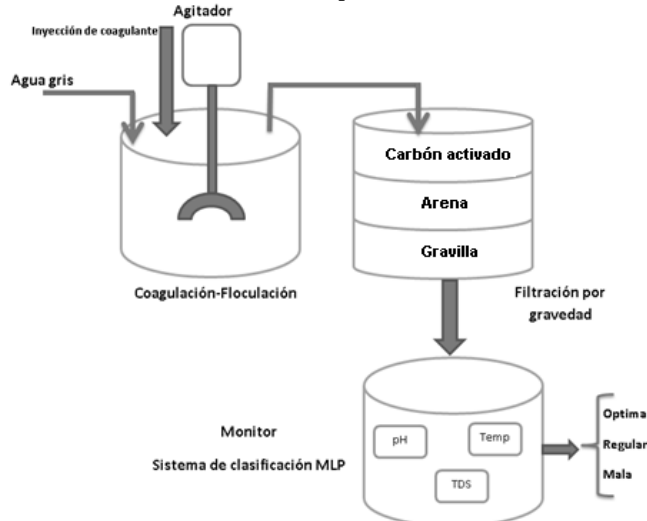


Figura 3. Diagrama a bloques del proyecto general

En la figura 4 se describe las variables que intervienen en el desarrollo del presente proyecto.

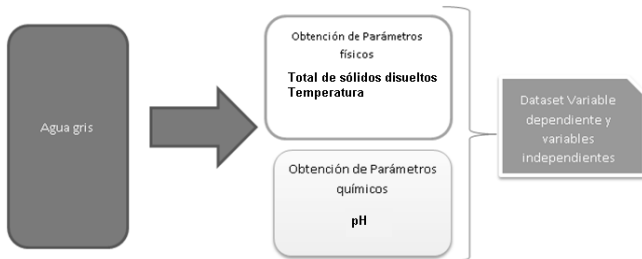


Figura 4. Variables para la obtención del modelo.

El objetivo de este proyecto es la generación de un modelo con redes neuronales para clasificar el agua residual gris a partir de un conjunto de datos previamente adquirido, el modelo obtenido se convierte a un modelo de Tensorflow Lite para sistemas embebidos en específico para una Raspberry donde se adquieren los valores de sólidos disueltos totales (SDT), pH y temperatura y con ello se implementa el modelo entrenado de clasificación para obtener la clasificación de agua residual gris. En la figura 5 se muestran los detalles para implementar el modelo Perceptrón multicapa en una Raspberry.

Cabe recalcar que por su empleo práctico y manejo de instrumentos los sólidos disueltos totales los llamamos en el presente documento TDS por sus siglas en inglés (Total dissolved solids).

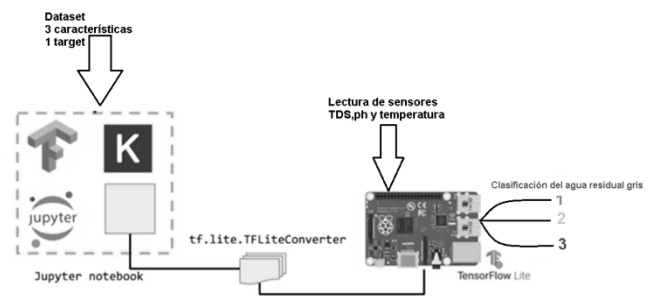


Figura 5. Conversión del modelo para implementar en Raspberry.

2.2 Sistema de adquisición de variables de temperatura, pH y total de sólidos disueltos.

En la figura 6 se muestran los sensores empleados, la sonda de pH 4502 C, que permite medir tanto valores de pH como de Temperatura y la sonda para medir los sólidos disueltos totales TDS V1.0



Figura 6. Sensor de pH –Temperatura y TDS

Los sensores cuentan con salida analógica por lo que para la conexión con la Raspberry se requieren un convertidor analógico-digital, para ello se empleó el convertidor ADS1115, éste cuenta con 4 ADC de 16 bits y se conecta con protocolo I2C hacia la Raspberry.

EL conjunto de datos cuenta con tres características (TDS, pH y temperatura) contiene 30000 muestras, de las cuales 5000 de clase 1 (optima), 10000 de clase 2 (regular) y 15000 de clase 3 (baja).

Al momento de desarrollar el presente trabajo se cuenta con un conjunto de datos con tres características TDS, pH y temperatura y un target que es la clasificación (1=Optima, 2=Regular, 3=Baja). En la tabla 1 se detallan los valores.

Tabla 1. Conjunto de datos para la obtención del modelo.

	tds	ph	temp	clasificacion
0	31	8.961454	13.437	1
1	31	8.973645	13.437	1
2	31	8.969073	13.437	1
3	31	8.963740	13.437	1
4	31	8.963740	13.437	1
5	31	8.963740	13.437	1
6	31	8.966026	13.437	1
7	31	8.960692	13.437	1
8	31	8.962978	13.437	1
9	31	8.961454	13.437	1

2.3 Sistema de clasificación con red neuronal multicapa tipo Perceptrón.

Para la implementación del sistema de clasificación empleando Redes neuronales se utilizó la plataforma de Jupyter que se base en formatos de cuaderno de programación y las librerías de ScikitLearn y TensorFlow. Enseguida se detallan los pasos para la obtención del sistema de clasificación.

2.3.1 Lectura del conjunto de dato, pre-procesamiento y visualización de datos

Una vez que se cuenta con el conjunto de datos que contiene los valores de pH, TDS y temperatura (ver figura 7), así como el valor de clase u objetivo, se emplea la librería pandas para realizar la lectura del conjunto de datos con extensión csv.



Figura 7. Lectura y carga del conjunto de datos.

A través del comando fillna de la librería de pandas se eliminan los datos vacíos que puedan existir dentro del conjunto de datos, generalmente algunas técnicas rellenan con ceros o unos dichos datos.

Se utiliza la librería de Seaborn para realizar gráficos representativos del conjunto de datos, a continuación se muestran gráficos representativos de las características de entrada contra los valores de clase.

En la figura 8 se muestra el gráfico de dispersión entre las variables de pH y sólidos disueltos totales junto con el valor de

clasificación, se puede interpretar que los valores de pH rondan los valores entre 7.5 y 12.5 y los valores de sólidos disueltos entre 30 y 500

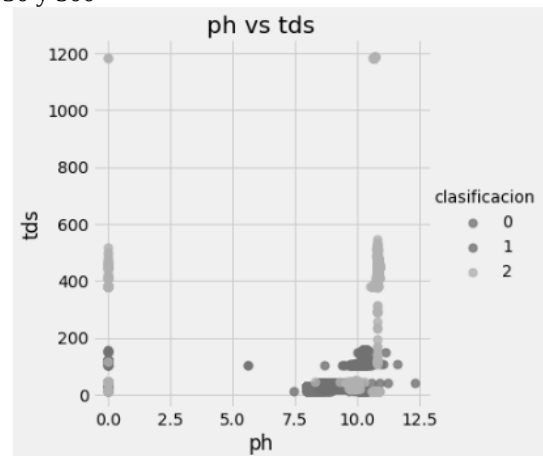


Figura 8. Gráfico de dispersión entre pH vs TDS

En la figura 9 se muestran los histogramas de valores de pH con respecto a la clase, se puede deducir que los valores de pH menores a 7 caen dentro de clase 0 (optima), los valores entre 7-10 caen dentro de clase 1 (regular) y los valores mayores a 10 caen dentro de la clase 2 (baja).

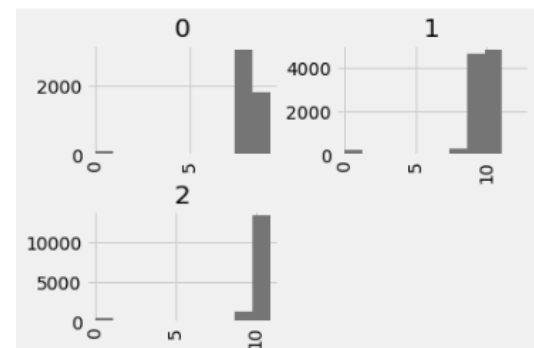


Figura 9. Histogramas de pH con respecto a la clase.

A partir del conjunto de datos original se procede a separar el conjunto de datos de entrenamiento y prueba para ello se emplean las herramientas de la librería Scikit Learn y otras utilidades del módulo Keras de TensorFlow.

2.3.2 Implementación de la red neuronal MLP

Se utiliza la librería de TensorFlow para generar el modelo de red neuronal tipo Perceptrón Multicapa MLP, la capa de entrada se define con 3 entradas y función de activación 'relu', se definen dos niveles de capas ocultas cada una con 500 neuronas y función de activación 'relu' y la capa de salida se define con 3 salidas con función de activación 'softmax' (ver figura 10).

La arquitectura del modelo consta de una capa de entrada, dos ocultas y una capa de salida, se elige esta arquitectura debido a la robustez de la red, donde solo se tienen tres clases en el vector de salida de la red. Además dicha red se implementa en un sistema embebido capaz de implementar el algoritmo de clasificación. En la evaluación y prueba se utilizaron distintos hiperparámetros: capas ocultas, la profundidad de las capas, la función de activación, hasta llegar al modelo de red propuesto.

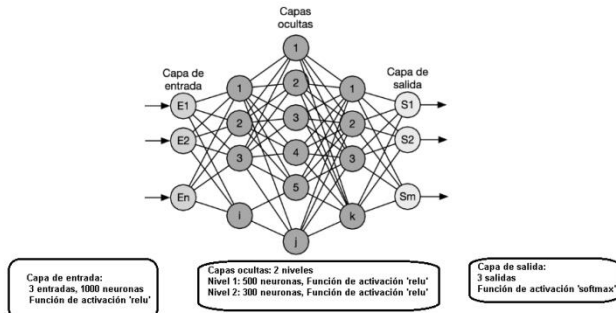


Figura 10. Modelo de red multicapa.

Una vez que se crea el modelo de red neuronal se procede a la compilación de la red neuronal donde los hiperparámetros relevantes son la función de pérdida y el optimizador, se usó Adam como algoritmo optimizador ya que es computacionalmente eficiente y tiene pocos requisitos de memoria. Como función de pérdida se utilizó entropía cruzada categórica. Tras la compilación del modelo se procede al entrenamiento para ello se emplea los datos entrenamiento y prueba.

Se obtienen los valores de precisión (score) para el modelo de aprendizaje con redes neuronales, los valores obtenidos una vez entrenado es del 99.48% para los datos de entrenamiento y un score de 99.87 para los datos de prueba (ver figura 11). En la siguiente figura se muestra el grafico con los valores de precisión obtenidos tanto con los datos de entrenamiento y los datos de prueba.

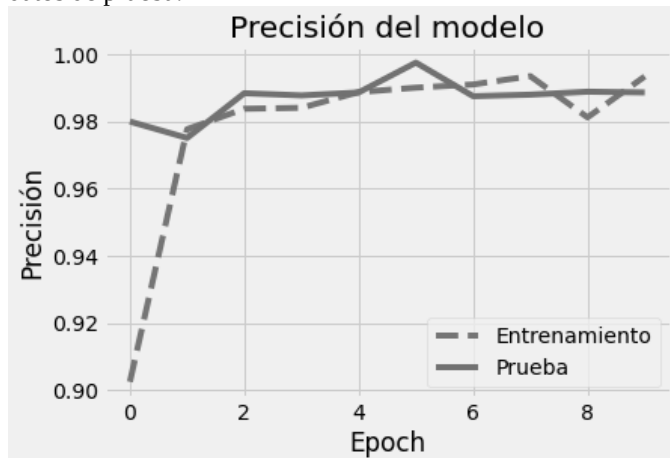


Figura 11. Precisión del modelo.

3. RESULTADOS

Al momento de realizar este artículo se han realizado diversas pruebas de validación e implementación del modelo en sistemas embebidos robustos como una Raspberry Pi.

Para medir la precisión del modelo en sistemas de clasificación se utiliza la matriz de confusión.

Para obtener la matriz de valores y graficar la matriz de confusión se emplea la librería Scikit learn, una vez que se tiene dicha matriz se procede a graficar empleando herramientas de la librería de Seaborn en específico el mapa de calor. En la figura 12 se muestra la matriz de confusión obtenida, en esta matriz se observa que de 1483 datos de clase 1, 1480 los clasifico dentro de la clase 1, mientras que 3 en clase 2. También de 3043 datos de clase 2, 3037 los clasifico dentro de la clase 2, mientras que 3 en clase 1 y 3 en clase 3. Además de 4474 datos de clase 3, 4471 los clasifico dentro de la clase 3, mientras que 3 en clase 2.



Figura 12. Matriz de confusión.

3.1 Conversión a modelo TensorFlow Lite

Una vez que se guarda el modelo en formato Keras H5, el modelo se puede convertir a un modelo de TensorFlow Lite más robusto y compatible con sistemas embebidos como Raspberry Pi entre otros. En la figura 13 se muestra los formatos de dicha conversión.



Figura 13. Conversión a modelo LITE.

4. CONCLUSIONES

La creación de nuevas herramientas y tecnología para el tratamiento de agua residual es un reto tanto para investigadores y centros de investigación, así como de los mandos de gobierno estatal y federal en México, ya que los niveles de los mantos acuíferos son bajos. Además los niveles de sequía extremos en México han propiciado la escasez del

vital líquido en diversos estados. Es por esto que el cuidado y manejo del agua debe incluirse en todas las agendas municipales y estatales y contar con planes bien definidos para su uso.

Desde un punto de vista personal como investigador y docente debemos enfocar nuestros proyectos a problemáticas reales como lo es el cuidado y manejo del agua, que en el marco contextual ubicado en la ciudad de Durango es un rubro de máximo interés debido a todas las problemáticas ya mencionadas como la sequía extrema entre otras.

Desde un punto de vista práctico el proyecto permite contar un instrumento doméstico para tratar el agua residual gris con el valor agregado del uso de la inteligencia artificial (IA) mediante redes neuronales que permite clasificar el uso que se le puede dar al agua tratada.

En lo que respecta a futuras líneas de investigación, el modelo propuesto de clasificación puede perfeccionarse en diversos sentidos, como lo es utilizar otro modelo de inteligencia artificial buscando que sea más robusto, otra manera, es implementar el algoritmo inteligente en otros sistemas embebidos capaces de soportar los motores de instrucción para modelos de inteligencia artificial.

5. REFERENCIAS

- [1] CONAGUA (2020) Aguas subterráneas Disponibilidad de Acuíferos. Consultado el 15 de Septiembre de 2023 https://sigagis.conagua.gob.mx/gas1/Edos_Acuiferos_18/durango/DR_1003.pdf
- [2] CONAGUA (2024) Monitor de sequía en México. Consultado el 5 de Junio de 2024 <https://smn.conagua.gob.mx/es/climatologia/monitor-de-sequia/monitor-de-sequia-en-mexico>
- [3] CONAGUA (2024) Monitor de presas en México. Consultado el 12 de Junio de 2024 <https://sinav30.conagua.gob.mx:8080/Presas/>
- [4] INEGI (2018) Usos del agua en México. Consultado el 10 de Septiembre de 2023 <https://cuentame.inegi.org.mx/territorio/agua/usuarios.aspx?tema=T>
- [5] Allen L. (2015) Manual de diseño para manejo de aguas grises. Consultado el 20 de Septiembre de 2023, <https://greywateraction.org/>
- [6] De Bazant J. (2009) Hacia un desarrollo urbano sustentable: problemas y criterios de selección. ISBN: 9786070505942
- [7] Calvo D., (2017), Clasificación de redes neuronales artificiales, Consultado el 21 de Agosto de 2023, <https://www.diegocalvo.es/clasificacion-de-redes-neuronales-artificiales/>
- [8] TensorFlow (2023), Perceptrones multicapa para el reconocimiento de dígitos con API centrales, Consultado el 21 de Agosto de 2023, https://www.tensorflow.org/guide/core/mlp_core